

Good government for small countries

The tendency for the world to be sprinkled with small countries has recently accelerated, and is disconcerting, but may yet offer a solution for British Ulster, where 26 people have been killed since the year began.

ARE we heading for a world of microstates? One perplexing accompaniment of the break-up of the Soviet empire is the emergence of would-be nation states with populations of five million or so, an order of magnitude less than those of even middle-sized powers such as Britain and France. Apart from the three Baltic states, there are at least Romanian Moldavia, Slovenia and Croatia (the pieces into which Yugoslavia has already broken); Slovakia hankers after separation from Czechoslovakia and Mr Boris Yeltsin has yet to come to grips with the centrifugal tendencies of the autonomous regions with which vast Russia is dotted (but he must know that something will have to be done about North Ossetia, in the Caucasus, which wants separation from Russia, but not union with Georgia to the south).

Fissiparous tendencies are more widespread. Belgium is already two separate states (see *Nature* 355, 579; 1991). Now, according to the pre-election British opinion polls, there is a resurgence of Scottish opinion in favour of separation from the United Kingdom (which opinion the opposition Labour party would mollify with an elected Scottish parliament). Unconvincingly, Welsh nationalists have taken up the cry. Where will it all end?

In further decentralization; there is no doubt of that. Events in the former Soviet Union have strengthened hankerings after democratic procedures and self-determination everywhere, while economic groupings such as the European Communities have made the option more credible: given a free market in goods and services, quite small parts of Europe may prefer to opt for formal self-determination in regional matters.

So why is the tendency towards microstatehood perplexing? It is not just that there are now more names to remember, and a greater variety of national aspirations for diplomats to take account of. More seriously, microstates can rarely be economically self-contained (unless they are strictly agrarian, as enlightened as Costa Rica or, otherwise, as technically sophisticated as Switzerland). Yet with the civil war in Nicaragua ended, even the handful of microstates in Central America might be drawn into a productive economic union (which the United States is best placed to foster).

The more substantial hazards are not economic. Usually, microstates can defend themselves only by alliances with others or with more powerful patrons — the recipe for the perfidious horse-trading that made eighteenth-

century Europe a nightmare. Worse still, because many of the emerging microstates are founded on an ethnic or cultural identity, their existence assaults the liberal doctrine that modern civility requires that people who think of themselves differently should nevertheless be able to live together productively.

The immediate danger is that newly independent microstates founded on supposed national aspirations will oppress the minorities they embody. What exactly are the arrangements in the Baltic states to safeguard the rights of their substantial Russian minorities, or in Croatia (as defined by the ceasefire line) for the protection of the Serbs who live there? What protection is there, throughout Central Europe, from anti-Semitism? Recent welcomes for emergent microstates as proofs that illiberal governments have fallen should have been more widely tempered by insistence that seemly rules should be in place — and that they will be followed. Otherwise, the immediate risk of unjust oppression will be followed by that of conflict on behalf of the oppressed (or supposedly oppressed) — Europe's 1930s nightmare. Luckily, there is now a framework within which such complaints can be raised: the Conference on Security and Cooperation in Europe (CSCE). It must be used more vigorously.

But there are, sadly, several microstates in which this remedy will not serve. One of them is British Ulster, where 26 people had been killed in the first six weeks of this year. (Others are Nagorno-Karabakh, the Armenian enclave in Azerbaijan, Indonesia's East Timor and Kashmir.) Ulster is not, of course, an independent state, but a part of the United Kingdom. The electoral majority, mostly Protestant by religious persuasion, want things to stay like that. Some others, variously described as nationalists or republicans, would prefer *anschluss* with the Republic of Ireland (whose constitution declares the aim of islandwide authority, but which also gives the Roman Catholic Church a special place).

Despite an agreement on consultation between London and Dublin (over the objections of the Protestant Unionists), things go from bad to worse; more people are killed, and business life is increasingly corrupted by protection rackets. Now that microstates are back in fashion, should not Ulster (with nearly two million people) join their ranks, but on the understanding that the British and Irish governments would share responsibility for keeping the peace so long as that is necessary, the next half-century or so? □



Europe's false star

The European budget for the next five years would spend too much on applied research, too little on other kinds.

SINCE the 1960s, Europe has been mesmerized by what was then called the "technology gap" between itself and the United States; now, there is Japan as well. For much of that time, European governments have looked to the European Commission to put things right. A French plan devised by the influential Pierre Aigran in the early 1970s would have made the Commission responsible for the development of a modern telecommunications network, for example. Since then, most of what is spent under the European Communities' (EC) research programme has gone to support supposedly "pre-competitive" applied research in support of advanced technology. Now, the Commission's budget for the next five years would spend even more. The ambition is misconceived.

Not that the EC is irrelevant to the technology gap. On the contrary, one of the chief reasons why it exists (and why there is now a throng of would-be new members) is that it will create a market for goods and services large and flexible enough to support industrial companies able to break new ground in advanced technology. By that means, Europe can reasonably hope to secure a substantial share of world trade in the future. To be sure, it will not happen overnight, or even the instant at which the arrangements promised by the Treaty of Rome are made a reality at the end of this year. But it will happen, and that should be sufficient.

So why does the Commission, and the EC which the Commission serves, keep hankering after artificial stimulants of advanced technology? It is not that the schemes already mounted, for example the *Esprit* programme centred on the interface between electronic communications and computers, have been resounding stimulants of saleable technology or even distinguished essays in applied research. But it seems that the Commission must be seen to be active in the field, if only to pacify the governments of France and Italy. In the past few weeks Mmme Edith Cresson, the prime minister of France, has been especially active in advocating that Europe should adopt some kind of industrial policy, apparently oblivious of past French failures in the same field.

Yet even cosmetic research programmes can be damaging. To the extent that European companies regard participation in Commission projects seriously — they usually pay half the cost — they may be tempted to forgo the thorough integration of research and development with their manufacturing operations that is the hallmark of success elsewhere. But the standard terms of Brussels projects, requiring collaboration between companies and research groups from more than one member state, also blunt the edge of the competition on which economic success depends. And the Commission's preoccupation with applied research diverts it from what should be its central task in the management of research, that of foster-

ing the education in research of as many young European men and women as it can find. The Commission should be spending more on basic, not applied, research. To its shame, it has been cutting back. □

Indian rope trick

The reinstatement of Professor V. J. Gupta as a member of the faculty at Chandigarh is a surprise, to say the least of it.

ONE of the oldest tricks in the world has a person throwing one end of a rope vertically upwards and basking in the applause of his audience when it does not fall back towards the Earth. To emphasize that gravity has been defied, the performer even climbs the rope, waving to his audience as he goes. Properly done, it is an excellent trick. The Panjab University of Chandigarh, India, has now performed an academic version of it by reinstating Professor V. J. Gupta on its faculty, from which he was suspended just a year ago (see *Nature* **349**, 645; 1991). Gupta is the palaeontologist accused of having published a long series of research articles based on fossils whose provenance seems rarely to have been the Himalayan strata and locations to which Gupta attributed them.

Chandigarh's indulgence of Gupta is a kind of rope trick in that it defies the admittedly unwritten laws that usually apply when people are accused of publishing fraudulent data. Of course, an accusation of fraud, even when based on such detailed and specific evidence as that adduced by John Talent of Macquarie University (*Nature* **338**, 613–615; 1989) is not a proof of guilt. What the university should have done was to set up an independent inquiry into the Gupta case, made the conclusions public and then moved to a determination of the case. There has indeed been an inquiry and two field expeditions, but the reports have not been made public. Instead, they have been considered privately by a retired high court judge together with Gupta's defence against his accusers. Meanwhile, Gupta has remained physically in charge of his research material during his suspension.

This may satisfy the Panjab University, but is unlikely to cut much ice elsewhere. (Gupta's reinstatement raises the strange prospect that colleagues at Chandigarh who courageously voiced their doubts may themselves now be subjected to official inquiry.) One feature of the case against the authenticity of Gupta's fossils is that it has provoked the appearance of further evidence; thus researchers at Aberystwyth now claim a fossil published by Gupta and attributed to the Himalayas is identical to one missing from their university's collection. While such allegations persist, there are few outside India who will in future collaborate with Gupta. Even in India, and despite the patronage a powerful professor can bestow, helpers may become scarce. Chandigarh may shrug off the consequences, but other Indian universities and research institutes manfully (and often successfully) seeking international reputations and collaborations will not be so sanguine. □

Researchers ask for help to save key biopesticide

- Crop pests already showing resistance to *Bt*
- Report calls for industry, government curbs

Washington

AGRICULTURAL researchers want the US government and industry to adopt a more careful use of the biopesticide *bacillus thuringiensis* (*Bt*) to reduce the chances that crop pests will develop resistance to its toxicity.

In a report to be delivered later this month to officials from the US Department of Agriculture (USDA), government and academic researchers call for precautions and potential regulations that would prevent insects from becoming immune to *Bt*. At stake is the viability of one of the most promising biopesticide ever developed.

Forty years ago *Bt* was just a bacteria found in ordinary dirt. Then, when researchers discovered that *Bt* produces toxic proteins that kill dozens of common insects (but are harmless to humans), it became a biopesticide to spray on crops. Over the last decade, however, as researchers discovered how to engineer the DNA of many plants so that they can express *Bt* toxins themselves, the simple bacteria has triggered a multi-million dollar research effort in agricultural industry to produce self-protecting crops that do not require chemicals to fight off pests.

Now, on the verge of a revolution in agriculture, researchers are afraid that *Bt*'s miracles could be coming to an end. Several US teams have managed to create *Bt*-resistant pests in the laboratory in as little as 12 insect generations. And if it can be done in the laboratory, it can eventually happen in the real world, they argue, especially when *Bt*-engineered plants reach widespread use.

Insects that are susceptible to *Bt* toxins are "a genetic resource" says Mark Whalon of the Michigan State University. "If we're not careful, we're going to lose them".

Last month, at a meeting sponsored by the USDA, government and academic researchers hammered out a strategy they hope can save *Bt*. Their report is expected to contain 124 suggestions for research on *Bt* toxicity, monitoring programmes and precautions to avoid insect resistance.

Chief among the recommendations is an independent national advisory body that would serve as a *Bt* guardian. Reporting to the USDA, the group would recommend research on the biopesticide and monitor its use. Farmers will have to be careful not to rely on *Bt* alone, says Whalon, one of the authors of the report. Instead, he says, they should try to "keep the insects

off balance" by mixing *Bt* with some traditional chemical pesticides, natural predators and other biotoxins.

Researchers also hope to encourage farmers to plant non-*Bt* plants alongside their engineered cousins that will serve as a haven (and genetic reservoir) for *Bt*-susceptible pests. Varying the percentage of *Bt*-expressing plants from zero to 50 per cent in a single field should also help to lower the percentage of pests with resistance by allowing susceptible insects (which usually have some other reproductive advantage over resistant insects) to



recover and drive out the competition.

Whalon has been able to breed Colorado potato beetles in the laboratory that show a 200-fold increase in *Bt*-resistance in 17 generations. He says that a comparable evolution would take just five years in nature. Whalon also bred his insects to be extremely susceptible to conventional chemical pesticides, so they could be quickly eradicated if they ever escaped. But naturally evolved resistant insects, in all likelihood, would not be so conveniently crippled.

Much of the *Bt* resistance that has appeared in the field is thought to be due to massive crop spraying of the *Bt* toxins themselves. The spraying has exposed insects to high doses and accelerated their rate of evolution into increasingly resistant breeds. Reports from at least four countries, including the United States, suggests that overuse of *Bt* by farmers has already produced some insects twice as resistant as those of just a few generations ago.

Officials at Monsanto Chemical Corporation, which is running one of the largest programmes to insert *Bt* resistance into crop plant DNA, say that its engineered plants are less likely to produce resistant pests. They explain that the toxins will come from inside the plants rather than from the outside, where they can wash away. That difference reduces the quantity of *Bt* that is released into the environment. In addition, the presence of *Bt* in the plant makes the insects vulnerable as soon as they begin to eat, at a point in their development when they are most susceptible.

Monsanto, too, is calling for a crop-management "regime" for *Bt*-engineered plants. If such precautions are not adopted, company officials fear that their star performers will have lost their lustre by the time they reach the market in 1995.

Calling *Bt* susceptibility a "national resource", the researchers who met at the federal conference want the government to issue regulations on the use of the biopesticide. To date, however, federal agencies have not shown much interest. Unless the agencies have a change of heart, the best the *Bt* sentries can hope for are federal guidelines and help from companies such as Monsanto in encouraging farmers to use the engineered plants with care. But if the experience with sprayed *Bt* is an guide, that will be an uphill battle.

Christopher Anderson

JAPANESE SATELLITE

Earth monitor fails

Tokyo

JAPAN'S first earth resources satellite, *Fuyo-1*, which was launched last week by the National Space Development Agency (NASDA), has run into serious trouble and may not be able to fulfil its main mission to search for mineral and oil resources around the world with a new high-technology radar system.

Shortly after launch, the 12-metre antenna of the Synthetic Aperture Radar (SAR) — the satellite's key sensor — failed to open fully on command from the ground. The reason for the jamming remains unknown.

The other optical sensors on the satellite are functioning normally, but they cannot see through cloud cover. The SAR is an active sensor that bounces microwaves off the Earth's surface and can determine surface characteristics, such as undulations, with a resolution of 18 metres regardless of cloud cover.

Fuyo-1 is the largest satellite ever launched by Japan. NASDA engineers have not yet given up hope; they say there are cases where a satellite's jammed antenna has opened, without permanent damage, as long as three months after launch.

David Swinbanks

Gaps loom in satellite data

London

CURRENT plans to monitor the global climate and stratospheric ozone depletion by satellite leave yawning gaps in the collection of key datasets, according to leading atmospheric scientists. As a result, the world's governments may be 'flying blind' — deprived of a continuous record of the environmental changes that they are striving to control — as they struggle to build and enforce international treaties to limit global change.

Officials on both sides of the Atlantic are aware of the problem. Last week, leading atmospheric scientists descended on Washington to argue the case for a series of new environmental monitoring satellites, to be launched in the mid-1990s. British officials are leading a drive to improve the international coordination of satellite launches, to prevent existing data gaps from widening still further.

But the gaps cannot be removed altogether without new satellites. And with most of the world's space agencies feeling a financial squeeze, some interruptions in the flow of data seem inevitable.

David Fisk, chief scientist at the UK Department of the Environment, says that a large gap in the satellite monitoring of stratospheric ozone depletion is likely to open in 1995, when the US-led Upper Atmosphere Research Satellite (UARS) stops functioning. Researchers had hoped that UARS would be followed by a similar German satellite called ATMOS. But Arndt Langner, head of Earth observation at DARA, the German space agency, says that Germany has "nearly given up" on its plans for ATMOS because of budget problems. Instead, he says, the agency is concentrating on salvaging some of the ATMOS instruments for later missions planned by the European Space Agency.

This leaves a three-year gap until the earliest possible launch of a similar satellite, as part of the international Earth Observing System (EOS). But the latest plan for EOS from the US National Aeronautics and Space Administration (NASA) would delay further the first launch of instruments to monitor changes in stratospheric ozone. NASA is making ready plans that call for the launch of the EOS atmospheric chemistry experiments in 2000 and 2002, rather than the original timetable of 1998.

Mark Schoeberl, an atmospheric physicist from NASA's Goddard Space Flight Center in Maryland, says that the gaps in data coverage are a serious concern. Other satellites will be carrying Total Ozone Mapping Spectrometers, which measure the total amount of ozone in the column of atmosphere beneath a satellite, he says. But Schoeberl says it is also important to record in detail the chemical changes oc-

curing in the stratosphere — data that ATMOS and the instruments originally due for launch in 1998, as part of EOS, would have provided. "Without a little more detail, it's very difficult to determine what's going on," he says.

The situation for climate monitoring is in many ways worse, largely due to the sheer complexity of the system that climatologists wish to monitor. Again, EOS will provide valuable data, but many researchers argue that there is also a need for a suite of 'cheap and cheerful' satellites to provide baseline data. Most of the instruments planned for EOS are designed for "diagnostic" studies into processes affecting the climate, rather than the basic job of climate monitoring.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

What will fill the void when UARS dies?

One major component of this less-glamorous task is to produce a continuous record of the Earth's radiation balance. But again, there are gaps in data coverage. There is now no instrument in orbit to monitor in detail the influence of clouds on the radiation balance, after the 1990 failure of a scanner on the US Earth Radiation Budget Satellite (ERBS).

Bruce Barkstrom, from NASA's Langley Research Center in Hampton, Virginia, says that a similar scanner may fly in 1995 on a joint mission between France and Russia. But even if the impoverished Russian space programme is able to launch the satellite as planned — which is far from certain — Barkstrom says the instrument is of a subtly different design. The difference, he says, will make it difficult to distinguish real changes in the Earth's radiation balance from differences in the behaviour of the two instruments.

To address the problem, Barkstrom is

proposing that the United States should launch a small satellite in 1995, carrying an instrument similar to that which flew on ERBS. He presented his plan in Washington last week to the federal Committee of Earth and Environmental Sciences, a body comprising officials from all the US government agencies interested in environmental research. The committee was also briefed by James Hansen, from NASA's Goddard Institute for Space Studies in New York, on a proposal to launch a sequence of small climate monitoring satellites, or Climsats.

Hansen's idea is to monitor a range of climate parameters using a small number of tried-and-tested instruments. The first Climsat could be launched for as little as \$80 million, and replacement copies could be launched every five years at \$40 million each.

But NASA has little money to put into new satellite programmes. Shelby Tilford, NASA's director of Earth science, accepts that the agency's current climate monitoring programme "doesn't do everything. The budget doesn't permit that," he says. But he says that NASA is trying to persuade the Department of Energy — which may have some money available in its 1993 budget — to support Barkstrom's proposal for a small radiation satellite.

With little money to spend on new satellites themselves, British officials are trying to prevent any widening of the gaps in data coverage that are now opening up by improving coordination among the various nations that are planning to fly environmental satellites. John Major, the UK prime minister, has described the British concerns in a letter to the other leaders of the G-7 nations. Major suggested that the problem could be addressed at a meeting of the Committee on Earth Observations Satellites (CEOS), to be held in London in late April.

The committee was originally an offshoot of the G-7, but its membership now includes most of the countries with an interest in remote-sensing satellites. The British would like this year's annual CEOS meeting to follow a meeting of the principal users of environmental monitoring satellite data. Once the scientists have outlined their requirements, officials from the major satellite agencies could discuss how to juggle their launch plans to provide the best possible data coverage.

The British plan seems to have been well received, at least in the United States. "I wouldn't like to characterize the current level of coordination as inadequate," says Peter Backlund, NASA's assistant director for Earth science, "but you can always do better".

Many environmental scientists are more sceptical, however. They point out that no amount of international coordination can replace satellites that have been delayed or cancelled.

Peter Aldhous

FBI gives in on genetics

Washington

THE US Federal Bureau of Investigation (FBI) is responding to criticism that its DNA 'fingerprinting' techniques vastly overstate their reliability by sending researchers on a worldwide hunt for better data on the genetics of ethnic subpopulations.

Introduced in the late 1980s, DNA fingerprinting has been both hailed as a certain method of identifying a suspect and attacked by critics for producing false matches and mistaken convictions. Recently, that debate has focused on the technique's assumptions regarding population genetics.

One postulate is that three or more genetic markers can identify an individual with near-certainty, with the chances of random matches given as high as one in a trillion in some court cases. In determining the odds, the FBI and private forensic laboratories have usually compared samples with databases of either Caucasian, African-American or Hispanic genetic samples.

But FBI researchers have started to re-examine those assumptions. They are now working with forensic scientists in Europe, Australia and South America to examine DNA databases of ethnic subgroups to see if the world is really as heterogeneous as they have claimed. FBI officials also plan to change their policies so that DNA samples are compared, whenever possible, with profiles obtained from nearby regions rather than the usual large race-based pools for the population at large. According to the newsletter *Bio-technology Newswatch*, FBI director William Sessions informed Congress last month that the agency would in the future use samples from more than 50 US geographic areas.

John Hicks, assistant director of the FBI's laboratory division, says that FBI staff scientists will travel to Denmark, Germany, Italy, France and Spain to collect data from forensics DNA laboratories in those countries. Cellmark Diagnostics, a UK-based forensic laboratory, will collect data from other countries, including aboriginal populations in Australia. A typical database of this sort contains DNA fingerprints of between 200 and 300 individuals. Although the FBI hopes to obtain some data from Indian populations in South America, Hicks says, Black populations will not be included in the survey because comparable data is not available.

The FBI's measure is in part a response to the criticism of two population geneticists — Richard Lewontin of Harvard University and Daniel Hartl of the Washington University School of Medicine in St Louis — who argued last December in *Science* (254, 1745; 1991) that the FBI's

methods of determining chances of coincidental matches may overstate the effectiveness of DNA fingerprinting by orders of magnitude. The fact that many ethnic subgroups are heavily concentrated in certain parts of the United States, Lewontin and Hartl claim, could make the chance of finding a given genetic combination much more likely among those groups. Certain genetic combinations may be more common in Italian neighbourhoods, for example, than in the Caucasian population at large, they say.

Lewontin and Hartl proposed that researchers should study DNA samples from 10 to 15 ethnic groups in Europe to determine just how often certain genetic markers show up in ethnic populations that are relatively homogeneous. If the markers turn out to be relatively common in these ethnic groups, they argue, then forensic researchers should expect clusters of the groups in the United States also to show genetic similarity. That would increase the odds against a coincidental match of DNA samples among these groups.

The FBI says that its assumptions are "reasonable and conservative estimates" that compensate for the possibility of ethnic clustering. But agency officials have nevertheless decided to do much of what Lewontin and Hartl recommended.

When examining the new data, Hicks says, "we want to see if there is any great divergence from the kind of distribution of [DNA] alleles we see in the United States. We've assumed that there isn't."

In a 24 January letter to Representative Don Edwards (Democrat, California), chairman of the House Judiciary Committee's subcommittee on civil and constitutional rights, Hicks described the changes and said that the agency hopes to have the European data collected and analysed by this summer. Edwards plans to hold a hearing on the issue later this year, although no date has been set.

Hartl, although agreeing that the new FBI studies are the type he and Lewontin had called for, says that they are still too "sketchy" to provide conclusive proof about the need to use ethnic subpopulation databases in DNA fingerprinting cases. He points out that incompatible laboratory standards in the various countries could make valid comparisons impossible.

Next month the National Research Council of the National Academy of Science is expected to release a long-awaited report on DNA fingerprinting that will include a chapter on the increasingly controversial population genetics issues. But academy officials say that the report could be delayed because, like nearly everything else in this field, it has become embroiled in politics.

Christopher Anderson

Funding scheme dies

Sydney

THE Australian government has backed away from a plan to force pension funds to pour large sums into the commercialization of research after receiving strong criticism from academic and industrial groups.

The critics, in two recent reports, have cited poor results from previous government attempts to encourage venture capital and the lacklustre record of existing Australian industry in commercializing research. Instead of simply boosting the supply of money available for investment, the academic panel suggested, the government should pursue an industrial policy that encourages competition and reduces tariff protection as well as stimulating cooperative efforts by high-technology companies in selected fields.

A top-level task force of businessmen and scientists started the current debate in December with a report to Ross Free, the government's science minister, recommending that pension funds be encouraged to invest one per cent of their assets in venture and development capital projects. That contribution would be mandatory for funds with a significant level of overseas investments. Building upon the government's decision to set up a universal pension fund to be administered by the private sector, the one per cent requirement would eventually make available hundreds of millions of dollars for research and hopefully invigorate the country's sagging economy.

That idea was immediately criticized, however, by an independent government body, the Industry Commission. The commission issued a report concluding that an earlier government attempts to stimulate the supply of venture capital through special tax breaks had failed, with the money either being lost or yielding disappointing returns.

Last month, a report by the Centre for Technology & Social Change at the University of Wollongong said that the task force has concentrated too much on increasing the supply of research and venture capital. They called the government's proposal an example of "uphill string pushing".

Instead, the report said, the government should try to improve industrial demand for research by encouraging the formation of industry clusters in close geographic proximity.

Earlier this month the Industry Commission issued a second report that further questioned the government's efforts. The report said that a plan to encourage government departments to buy local, high-technology projects had been a waste of money. Some A\$25 million spent by the programme in four years, it said, had gone to projects that would have gone ahead anyway.

Mark Lawson

NIH merger to shorten review

Washington

THE Recombinant DNA Advisory Committee (RAC) of the US National Institutes of Health (NIH) voted unanimously last week to disband its human gene therapy subcommittee. The move will reduce, from six to five, the number of regulatory hoops researchers must jump through to obtain approval of federally funded experiments. It also resolves a communications gap between the two groups, which had some overlapping membership but which occasionally reached different conclusions about whether a specific experiment should be allowed.

RAC was set up by the NIH in 1975 to provide oversight for federally funded experiments involving the use of recombinant DNA technology. Before long, however, it became clear that many types of experiments, including the deliberate release into the environment of genetically engineered organisms, posed minimal risk to human health and the environment. Now, 17 years later, many of RAC's earlier responsibilities have either been deregulated or come under the aegis of other government agencies or local review boards.

In 1984 the RAC established what is now called the human gene therapy subcommittee to address the specific technical and scientific issues raised by experiments involving gene transfer and gene therapy in humans. It is because RAC and its gene subcommittee now perform much the same role that LeRoy Walters, a bioethicist from Georgetown University and chair of the gene subcommittee, thinks it is a "logical time to make this step".

By streamlining the process into a single review at this level, RAC hopes to become faster and more efficient and end the poor communication that existed be-

tween the two panels. With no guarantee that subcommittee members will be made full voting members of RAC, however, some members are worried that the parent committee could find itself lacking the expertise necessary to review experiments of this kind.

Researchers are also frustrated by the amount of duplication within the two-tiered system. Even though six members of the subcommittee also hold seats on the RAC, the lack of communication between the two committees has made it necessary for researchers to present the same data to both the gene subcommittee and the RAC and then to undergo two quite separate reviews. Martin Gellert, a laboratory chief in the metabolic enzymes section of the National Institute of Diabetes and Digestive and Kidney Diseases within the NIH and a member of both committees, admits the system broke down because there was "no mechanism for transmitting detailed discussions of the gene therapy subcommittee to the RAC".

Although the RAC has recommended that members of the subcommittee who are not also members of the parent committee should be appointed to the RAC, the final decision is out of RAC's hands. All nominations must be approved at the secretarial level of the department of Health and Human Services. Gellert believes that in the interim this could leave RAC short on technical and scientific know-how as there was a higher representation of scientists with knowledge of human gene therapy on the subcommittee. "Right now, the RAC is a little bit short on that kind of expertise", he says.

To compensate for the lack of preliminary review, RAC has agreed to increase the number of annual meetings from three to as many as six. **Diane Gershon**

ANIMAL EXPERIMENTATION

Swiss public backs research

London

THE Swiss electorate has rejected an initiative that pharmaceutical companies and many academic researchers claimed would paralyse research in Switzerland involving animals (see *Nature* 355, 575; 13 February 1992).

Paul Walter, from the University of Basel and president of the pro-research lobby group Arbeitskreis für Gesundheit und Forschung, says that the 56 per cent to 44 per cent vote was "a larger majority than we expected". In only four of the 26 Swiss regions, or cantons, did the initiative win majority support.

Walter believes that a large number of previously undecided voters were convinced by the arguments put forward by

the pro-research lobby in the build-up to the referendum on 16 February. Walter's group, universities and the pharmaceutical industry had argued that a clause giving animal welfare groups the right to challenge proposed experiments in the courts could halt most Swiss animal research.

But Hans-Peter Haering, secretary-general of Schweizer Tierschutz, the Swiss Animal Protection League, accuses his opponents of misrepresenting the group's initiative. "We didn't want to forbid all animal experiments," he says. Haering says that the Animal Protection League will focus its attention on tightening animal welfare legislation at the cantonal level. **Peter Aldhous**

Airlines vs activists

Washington

NORTHWEST Airlines, declaring that it "had no way to judge the merits of [animal rights] arguments", last week resumed shipments of laboratory animals after suspending them last fall because of protests from animal activists. The decision ends the company's self-appointed but naive attempt to become an arbiter in the bitter controversy over the use of animals in research.

Jon Austin, a spokesman for the company, which is based in Minneapolis, Minnesota, says that officials debated the issue for several weeks and found that the arguments "were equally compelling on both sides". They decided that the subject should be resolved by governmental bodies, not private companies. In the meantime, Austin says, Northwest is reverting to its original policy of accepting shipments that are "safe and legal".

In November the company decided to suspend its delivery of some 160 beagles to a Swiss laboratory, where they were to be used in toxicity experiments. Animal rights groups, including Friends of Animals, took credit for the action, which they said was part of a campaign to eliminate the use of animals in research. Austin had announced then that Northwest was "going to try to figure out what the socially responsible thing to do is."

The temporary ban triggered a letter-writing campaign from researchers around the United States, including several scientists at the Mayo Clinic in nearby Rochester, Minnesota. The airline received some 500 letters and phone calls on the issue, divided equally between supporters and opponents of animal research.

One surprising participant in the debate was a former presidential candidate, Walter Mondale, who sits on the board of directors of both the Mayo Clinic and Northwest. Ross Corson, a spokesman for Mondale, confirmed that Mondale conveyed the feelings of Mayo researchers to his fellow board members at Northwest. "He asked them to review their policy", says Corson.

The airline was handicapped by its lack of knowledge, according to Austin. "We would hear from somebody who was a member of the National Academy of Sciences, in support of animal research, and then we would hear from somebody else, with equally impressive credentials, arguing the opposite view," he says. "And we were stuck in the middle, lacking the credentials to make a decision."

Austin says that the decision to resume shipments does not mean that Northwest supports the use of animals in research, only that "we decided our responsibility was to follow the law, as it is currently written."

Jeffrey Mervis

Glaxo benches top manager

London

WHAT should a successful research organization do when one of its leading research managers decides he would like to return to the laboratory bench? For Glaxo, the answer is to spend £16 million to move the manager, complete with research team, to a laboratory in one of the country's best university pharmacology departments.

Not every company is able to make such a decision. But it is a bit easier when one is the world's second largest pharmaceutical company, and one's manager once discovered a drug that is on the brink of earning the company countless millions of pounds.



Glaxo's Pat Humphrey prepares for encore.

Under an agreement signed this week, Patrick Humphrey, director of pharmacology with the UK company Glaxo, will set up shop this spring in an empty laboratory at the University of Cambridge. The company hopes that the site will eventually be home to some 35 researchers, busy at work on a range of molecules including somatostatin, a peptide found in many tissues that is involved in a multitude of physiological processes.

The decision to send Humphrey off to a new laboratory represents more than just an expansive gesture to reward the goose that laid the golden egg, however. Richard Sykes, Glaxo's director of research and development, says that the company wants to be sure that its scientists, many of whom found themselves thrust into management during Glaxo's explosive growth over the past two decades, are working on what they do best.

In the past, says Sykes, "we have taken scientists, tried to train them as managers, and probably ruined their careers." Under the company's new policy, some small research teams will be freed from the need to report regularly to one of the company's research management committees. Instead, they can devote their energies to such key projects as the role of the bacterium *Helicobacter pylori* in peptic acid disease. In addition, Sykes says that the company wants to dispel the notion that a move into management is the only way to achieve rapid promotion.

Other British pharmaceutical companies believe they have already found the

right balance between promotion opportunities for managers and for researchers. "It's not a particular problem for us," says Keith Bradford, research and development personnel director with SmithKline Beecham. Norman Elmore, a research manager with ICI Pharmaceuticals, says that his company has comparable promotion ladders for scientists who choose to stay in research and those who move into management.

Maybe so. But as Humphrey points out, those companies have not had to contend with the rapid growth that Glaxo has experienced. "I think it's remarkable that Glaxo has managed to grow at the rate that it did without losing its way," he says.

Glaxo's growth has been formidable. In 1972, when Humphrey joined the company, Glaxo spent a mere £7 million on research and development; today, the figure is more than £450 million. On the strength of its research, Glaxo has been transformed from a nondescript food and drugs outfit into a company selling more than \$6,000 million worth of pharmaceutical products each year. That is a figure bettered only by its US rival, Merck and Co.

Almost 50 per cent of these sales come from a single product, the ulcer drug Zantac, or ranitidine. The drug that Glaxo hopes will continue this success — Imigran, or sumatriptan — was based on Humphrey's research into the neurotransmitter serotonin. Imigran, which is used to treat migraines, has already been licensed in several European countries, and is expected shortly to win approval from the US Food and Drug Administration.

Humphrey's new institute in Cambridge is designed to provide future generations of big-earning drugs for Glaxo. Together with a small team of Glaxo researchers (several of them also research managers returning to the laboratory), Humphrey will move in April to Cambridge from Glaxo's research site at Ware, in Hertfordshire.

Pharmaceutical companies have set up research groups in British universities in the past — SmithKline Beecham, for example, has a 12-person unit in the University of Oxford's clinical pharmacology department. But Humphrey is adamant that the Glaxo/Cambridge arrangement (which will earn the university some £5 million over the next 10 years) will be much more focused on drug discovery than its predecessors. And he is confident that his years spent in research management will not be a serious disadvantage.

"I've done it once, and I can do it again. I've never been far from the centre of research," says Humphrey.

Peter Aldhous

Patents, round two

Washington

FOLLOWING their controversial application last summer for patents on 347 partial human gene sequences, US government lawyers filed a patent last week for another 2,375 gene fragments.

The filing, which brings the government's total claim to about three per cent of the human genome, coincided with the publication in *Nature* of a paper by J. Craig Venter, a National Institutes of Health (NIH) geneticist who identified the sequences. In explaining the reasons for the second filing despite widespread criticism of its predecessor, NIH director Bernadine Healy said that her agency was taking the only prudent course, given the present uncertainties in patent law. She described it as an "interim policy" that would ensure that NIH do not lose any possible property rights should the sequences turn out to be patentable, while allowing researchers and the biotechnology industry to continue their debate.

Although the idea of patenting gene sequence in which the functions of the genes is unknown is still anathema to many researchers, the consensus has shifted towards the NIH since last year's protests (see *Nature* 353, 485; 1991). Given the possibility that the sequences could be found patentable even without known function, biotechnology industry associations and many patent experts now say that NIH had little choice but to file an application, lest a private company of foreign company do it first. Biotechnology companies would be more supportive, however, if they knew just how NIH would license the sequences if the patent is granted.

Reid Adler, director of the NIH's technology transfer office, says that the agency intends to consult with industry to develop such a policy. "We want to be very fair and very open about this," he says. He will be meeting with the Industrial Biotechnology Association this week.

Christopher Anderson

SPACE RESEARCH

NASA chief sacked

Washington

ADMIRAL Richard Truly, who as head of the National Aeronautics and Space Administration (NASA) helped the agency recover from the 1986 explosion of the space shuttle Challenger, was fired last week in a policy battle over the future direction of the agency. Although Truly's departure after nearly three years at the helm will go down in the books as a resignation, the former astronaut made no secret of the fact that he had been told to leave by President George Bush.

Truly will depart 1 April. No replacement has been named.

Christopher Anderson

Canadian drug firms boost R&D spending

Ottawa

CANADA'S patent drug manufacturers, buoyed by the federal government's endorsement last month of an international trade proposal that extends patent protection for pharmaceuticals, have announced new research and development investments that will total Can\$309 million over the next four years.

The government's move was a reward to the industry for more than doubling its R&D expenditures since 1987, the last time that patent protection was extended. It also is meant to prod the industry to continue that trend. "These measures will encourage increased R&D in Canada, providing high-paying, skilled jobs for the medical and scientific communities," asserted Michael Wilson, international trade minister, in announcing the government's support for the GATT proposal.

The trade proposal is part of the current Uruguay Round of the General Agreement on Tariffs and Trade (GATT), announced last December in Geneva. It is designed to set common standards on international property protection for 108

member countries. The drug patent portion of the agreement was aimed particularly at Canada, which among industrialized countries has a unique system of "compulsory licensing" for pharmaceuticals. This system allows generic drug companies to copy and market brand-name products almost immediately after development and testing by the innovator, provided they pay a fixed royalty fee for the duration of the patent. Such licences are almost always granted.

The system was intended to hold down drug prices through increased competition. But the effect was to give patent drug manufacturers here a shorter period of market exclusivity for their products than is found elsewhere, and thus to reduce their profits. Many such Canadian companies are subsidiaries of multinationals.

To deal with this problem, Canada passed a bill in 1987 that extended market exclusivity for drugs from 7 to 10 years. In return, the drug industry promised to increase research expenditures. Although this improved the innovators' competitive situation within the country, they were

still less well-off than those in the US, for example, with 14 years of exclusivity, and the Europeans, with 15 years.

The GATT proposal would abolish compulsory licensing, although all licenses issued before the December draft was proposed would continue to apply. In practice, Canadian companies would receive exclusivity for three years more than at present, and the proposal would not directly affect other industrial countries. At the same time, it would force drug manufacturers from developing countries to adopt similar patent protection, but only after a grace period of up to 10 years.

In return for the extension granted in 1987, Canadian patent drug companies pledged to increase over the next decade their research spending from 4.9 per cent of sales to 10 per cent. A 1991 study by the independent agency set up by the federal government to monitor their progress, the Patented Medicines Prices Review Board, shows that they had already almost reached that goal: In 1990 R&D spending was 9.2 per cent of sales, and the Pharmaceutical Manufacturers Association of Canada reported last October that its members were spending 10.1 per cent of sales.

The 1987 patent law change also created jobs. An association survey last July showed that the number of research jobs had grown by 48 per cent, from 925 to 1,327, in three years. Overall, the industry added 1,386 jobs between 1987 and 1990.

The industry has also kept its promise to keep drug prices reasonable. The government's monitoring board found that existing patented drug prices stayed below the country's consumer price index between 1987 and 1990, as well as below the guidelines set by the government.

David Spurgeon

EC RESEARCH

Delors asks for doubled budget

London

In the first shot of a year-long battle over the European Communities' (EC) research budget, European Commission president Jacques Delors last week proposed a five-year EC financial plan that would double the EC's spending on research and place greater emphasis on 'near-market' research. On 12 February Delors told the European Parliament in Strasbourg that, by 1997, he wants the EC to spend an extra 3,500 million ECU a year (about \$4,400 million) on research and training to increase economic competitiveness.

Officials in the EC member states are still waiting for a breakdown of Delors' spending plans. But the proposed 3,500 million ECU increase is expected to include more than 2,000 million ECU for research and development, doubling the EC's present annual spending on its Framework research programme. Several of the member states, in particular Britain, will balk at this figure, and as Britain retained its effective veto over increases in EC research spending at December's EC summit meeting in Maastricht (see *Nature* 355, 3; 2 January 1992), Delors' budget is unlikely to be accepted in full.

"Europe's competitive edge has been blunted," said Delors in explaining the need for his proposal. "Its research potential is being eroded." As a remedy, the Commission wants the EC to support more 'near market' industrial research —

projects aimed at specific industrial sectors and leading to marketable products. That approach is championed by Filippo Pandolfi, Commission vice-president in charge of research.

France and Italy are likely to support these plans. But the idea is expected to be criticized by Britain, with backing from Germany and the Netherlands. They argue that EC support for industrial research should be restricted to 'precompetitive' projects, designed to develop underlying new technologies.

The outcome of this debate may well depend on the result of the British general election, expected in April this year. Glyn Ford, leader of the British Labour Members of the European Parliament, expects that his party would be more receptive to Delors' plans for near-market research projects if Labour were to replace the current Conservative government. Ford, formerly a science policy researcher at the University of Manchester, believes the EC should not be restricted to 'precompetitive' research programmes.

"Our problem in Europe is competing with the United States and Japan in high technology," he says. Ford argues that the Japanese government spends a larger amount on near-market research than does the EC, despite its protestations to the contrary. "We need to look at what they do, rather than what they say they do," he says.

Peter Aldhous

UK FUND CREATED

HIV compensation

London

IN A sharp reversal of previous policy, the UK Department of Health announced on Monday of this week (17 February) that it is setting aside £12 million to compensate some 80 non-haemophiliacs who were infected with HIV by blood transfusions carried out in the mid-1980s under the National Health Service (NHS).

The government had previously refused to extend the compensation that has already been offered to the 1,200 British haemophiliacs who contracted HIV from transfused blood products. It argued that this would set a dangerous precedent for compensating people harmed by NHS treatment, where negligence could not be proved. Infected people with dependent children can expect to receive up to £80,500 in compensation.

Peter Aldhous

DNA profiling and the police

SIR — Peter Aldhous reports (*Nature* 355, 191; 16 January 1992) that the civil rights group Liberty is seeking to take London's Metropolitan Police to the European Court of Human Rights, questioning the legality of our database of DNA profiling results. While I take a positive view of Liberty's attempts to clarify the law, it may have chosen a case that does not serve its purpose well.

Contrary to the account in *Nature*, the DNA profile of Roy Williams (a man questioned during a 1988 murder inquiry and subsequently cleared after voluntarily submitting a DNA sample for analysis) has never been on our DNA database. The Metropolitan Police Forensic Science Laboratory (MPFSL) did not measure the DNA profile in question. (This was done before we had brought DNA profiling into operational use.) We were simply told by the commercial company that did the work that Williams' DNA profile did not match with our scene-of-crime sample.

The MPFSL has not sought to hide its DNA database. I have lectured on it at the Royal Society and in the House of Commons; we show it (without breaching confidentiality) to most of our visitors, and have described both the database and its contents in our submission to the Royal Commission on Criminal Justice.

The DNA results held in our database are those from unsolved cases, from convicted criminals and from those awaiting trial in cases for which they provided a comparison blood sample. The House of Commons Home Affairs Committee wants the scope to be greatly expanded, but this would require new legislation. Many people believe that DNA profiles should be recorded for anyone convicted of a crime of violence, or at least for all sex offenders. The MPFSL view is that this is for the public to debate and for parliament to decide. In the meantime we work within the law as we understand it.

For example, if we obtain the DNA profile of someone convicted of rape before the technique became available, but who happens to be cleared in the case for which he donated the DNA sample, we do not put his DNA information into the database. We simply retain the X-ray film and the normal entries in our laboratory notebooks. Even here there are difficulties. We may be criticized for not tearing up our notebooks when suspects are cleared by our work. But imagine the outcry if forensic science laboratories did start destroying entries in laboratory notebooks! Public laboratories should not have to work under conditions of such legal ambiguity. The public desire

for confidence in its criminal justice system points to a need for some of these areas of uncertainty to be sorted out.

In conclusion, I should like to mention that among all the troubles that have hit forensic science in recent times, the MPFSL has never been implicated — in spite of being the largest forensic science laboratory in Europe, with an annual caseload of about 17,000 cases. And, far from muddying the waters on DNA profiling, it was my personal initiative that brought an agreement on a standard DNA technology choice in Europe, without which national databases would not be possible. The MPFSL also offered to run its DNA database on behalf of the nation before any alternative was available.

The Metropolitan Police is contrite in one respect. Roy Williams was apparently not told immediately that he had been cleared when we passed that information on to the investigating officers. I should like to add my own apologies as a Metropolitan Police employee to the others that have been given in the past.

BRIAN SHEARD
(Director)

Metropolitan Police
Forensic Science Laboratory,
109 Lambeth Road,
London SE1 7LP, UK

IQ and diet

SIR — I should like to correct an impression created by your News story "IQ pills' land in trouble" (*Nature* 355, 285; 1992). The charges brought by trading standards officers refer to advertising claims. The Benton and Roberts paper (*The Lancet* i, 140; 1988) was careful to make no claims: having reported the finding that a vitamin/mineral supplement increased the nonverbal intelligence of a group of school children, it concluded that "the study must now be replicated, and . . . the underlying mechanism must be established".

I have recently published a study that reports increased intelligence scores when six-year-olds receive supplements (Benton and Cook *Personality and Individual Differences*, 12, 1151–1158; 1991). There are so far five studies that report positive findings in at least some of their sample. That two studies have reported negative findings can be explained easily by the obvious assumption that only some children respond, those with a poor diet.

DAVID BENTON

Department of Psychology,
University College,
Swansea SA2 8PP, UK

Gene therapy

SIR — I have no doubt about the enthusiasm of scientists at the US National Institutes of Health (NIH) for the future of gene therapy, but from *Nature* I expect a more critical evaluation than that in "Gene therapy on the move" (354, 429; 1991).

For example, NIH and corporate researchers may be able to avoid issues of prenatal use of DNA technology as a method of disease control through reproductive choice; or eschew epidemiological conclusions that methods of disease prevention are already mostly at hand in the form of behavioural and environmental change; or pretend that polygenic diseases such as cancer, heart and liver diseases are likely new targets of gene therapy; or confuse arteriosclerosis prevention by neonatal transfection of LDL genes — which will predictably affect fewer than 1 per cent of those afflicted with artery diseases — with prevention for large majorities by change in life habits; but your reporter should not fail to remind us of these oversights and simplifications.

And the quotation from the child with cystic fibrosis surely belonged to a charity 'telethon'. *Nature* reporters and NIH institute directors should all be required to have read Thomas McKeown's last book, *The Origins of Human Disease* (Blackwell, 1988), as background support against the ever-increasing hype from our biomedical establishment leaders who, recognizing that current gene therapy addresses less than 2 per cent of the total disease problem (monogenic diseases) in the Western world, are busy reconstructing disease aetiology altogether.

RICHARD C. STROHMAN
Department of Molecular & Cell Biology,
University of California,
Berkeley, California 94720, USA

Old argument

SIR — Zakaria Erzincioğlu (*Nature* 355, 195; 1992) considers it absurd that we refer to things or people as "ancient" or "old" rather than "early" or "young" when they happened or live a long time ago, for example, are trilobites *ancient* or merely *early*? Is Aristotle?

From our point of view which is generally the coordinate system we favour, trilobites and Aristotle are both; they were here early (earlier than we were), younger than we are, but from our habitual standpoint of looking back through time from where we happen to be at the moment, they are ancient and therefore older than we are.

RALPH ESTLING
The Old Parsonage, Dowlish Wake,
Ilminster, Somerset TA19 0NY, UK

Disallowing direct costs

SIR — The recent concern over indirect cost recovery at US universities is well-founded, but researchers should be warned there may be a more ominous shakeup if government auditors decide to take a stricter stand on the allowability of *direct* costs.

Direct line-item expenditures are approved by government agencies when research proposals are accepted for funding, and as such are implicitly considered allowable. But when the time comes actually to purchase equipment, pay salaries or procure services, researchers are constrained by a baroque set of government regulations (the Federal Acquisition Regulations) that are applied and often interpreted differently by each funding agency. Auditors can (and do) disallow certain direct costs when investigators or their institutions fail to

follow these rules.

For example, an overseas trip paid for with Department of Energy contract funds may be disallowed if the principal investigator fails to file a pre-trip approval request 90 days in advance and a trip report upon his or her return to the United States — even if the trip was included in the investigator's original proposal. The same trip can be made on a National Science Foundation grant without prior approval. Bizarre reporting requirements are often buried deep in the middle of grant and contract documents, and some agencies make only a reference to applicable regulations without including the actual text in their award notifications.

As a former grant and contract negotiator for a major US university, I can say with confidence that most researchers are not familiar with funding agency regulations, nor are most of their institutions capable of ensuring full compliance.

Arguably, it is the responsibility of research administrators to ensure that principal investigators (PIs) and laboratory directors to understand the conditions of their awards. Similarly, university accounting offices and internal audit departments are supposed to discover and correct unallowable expenditures before the feds do. But in practice, few universities are able to do any of this adequately. Some of the burden of compliance inevitably falls on the PI. PIs, however, have little incentive or time to slog through endless pages of bureaucratic legalese to find out what they can and cannot do. The result is that direct cost disallowances may well become a major concern as attention is focused ever more closely on the way universities spend government money.

BRADFORD J. HURLEY

313 Commonwealth Avenue,
Concord, Massachusetts 01742, USA

Correcting errors

SIR — Over the past decade, I have encountered serious difficulties in correcting errors in the work of others (see, for example, my report in *J. Chem. Educ.* **68**, 474; 1991). My observations follow. There are (1) authors who are unwilling to publish *Errata* to their proven errors; (2) (internal) barriers to publication of explicit comments on errors in the work of others; and even (3) some authors who, when asked to function as referees, obstruct publication of a paper correcting their own errors. In my opinion, *Ethical Guidelines to Publication of Chemical Research* (*Acc. Chem. Res.* **18**, 355; 1985) may not

sufficiently respond to these situations.

Because it is important that the published record should be as accurate as possible, I propose that the following rules should be adopted by the scientific community and by journals: (1) *errata* may be published by persons other than the original authors; (2) authors who prefer to correct serious errors in the published papers of others in the text of their own original articles should cite in the abstract the source of the errors being corrected; and (3) in the case of an article correcting errors, the author(s) of the errors should never be selected as anonymous referees (although, of course, their right of a possible response to a published criticism would still apply).

ZDENEK SLANINA*

Limburgs Universitair Centrum,
Department SBG,
B-3590 Diepenbeek, Belgium

*Permanent address: Czech Academy of Sciences, Prague, Czechoslovakia.

Continental drift

SIR — Alan Goodacre (*Nature* **354**, 261; 1991) brought to mind something I read a few weeks ago in an anthology of geographical and travel literature texts (*Der Deutsche in der Landschaft*, ed. R. Borchardt, Insel Taschenbuch 1218, p. 163, Frankfurt, 1989) which reads in translation:

... One cannot help thinking of a real and earlier coherence between Arabia and its neighbouring countries in the East and the West, tracing on the map the trapezoidal outline of this peninsula, which appears to be a fragment, separated by the narrow, parallel gaps of the Persian and Arabian Gulfs. Fragmented at a time when the levels of the crust, already cooled down or dried off and lifted up, floating on a more mobile substrate, were not yet fixed as they are today, and therefore whole systems of the landscape could break loose, while the gaps were filled with water, forming straits. ...

This was written by Carl Ritter in *Die Erdkunde im Verhältniss zur Natur und zur Geschichte des Menschen Wissen-schaften* Vol. 1, 62 (2nd edn, Berlin, 1832). The fact that this is the second edition makes it possible that the same thoughts had been set down and printed even earlier, but I do not have access at present to the first edition. As can be seen from the text, the row between Plutonist and Neptunist had not yet been resolved, which probably hampered the general acceptance of such ideas for a long time. It took more than 150 years to collect the evidence to support the theory of continental drift, which in turn requires it to explain the data.

W. EBERHARD FALCK

Ismes SpA, Via dei Crociferi 44,
I-00187 Roma, Italy

Fermi's mistake?

SIR — Your cover of 2 January 1992 may give the widest publicity yet to one of the minor mysteries of physics or physicists. The left oval shows Enrico Fermi at his blackboard, a scene also used in "The World of Enrico Fermi", a 1970 Project Physics film. Above his head appears in Fermi's clear script the definition for α , the "fine structure constant" of atomic physics, but it is clearly, shockingly wrong. Every physicist learns early that $\alpha = (e^2/hc)$ or about 1/137, and Fermi had lived with α for at least 35 years.

I was shocked on glimpsing this anomaly when I saw the film; it turns out that Norman F. Ramsey, professor of physics at Harvard University, had long ago noted this error and was equally puzzled by it.

The problem is that α is a dimensionless quantity, known by rote to every physicist, in form and value. It is the ratio of the orbital velocity of the electron in the hydrogen atom to the speed of light, for instance. The anomaly on the blackboard is not dimensionless, is bigger the smaller is the electric charge e , and it is difficult to believe that Fermi could have written it, even under the (by then not unfamiliar) stress of being photographed for posterity.

The most probable explanation is that Enrico Fermi, a great physicist, both in theory and experiment, and a man full of fun and humour, was having a little fun with the photographer.

RICHARD L. GARWIN

IBM Research Division,
Thomas J. Watson Research Center,
PO Box 218,
Yorktown Heights,
New York 10598-0218, USA

Hungarian science confronts its past

Christopher Anderson

Despite resistance, there is a revolution in progress among Hungary's research community. The signs are promising that the legacy of Soviet-style inefficiency will give way to a brighter future.

Budapest

IN Hungary, perhaps the best-prepared of all the Eastern European countries to throw off 40 years of Soviet-style science, some shackles are proving hard to lose. While the country as a whole wrestles with the job of simultaneously reforming an economy, a political system, an industrial base and a national attitude, Hungarian scientists are just trying to get their Academy to relax its stranglehold on the country's research institutions.

As an internal revolution, it is a curious thing to watch. Three years after the country itself abandoned the last vestiges of Communism, the scientific establishment is turning out to be surprisingly resistant to change.

Most of the academy's members were installed long before the political reforms, and everyone knows which were appointed because they were members of the Communist Party. The academy controls its institutions largely as it did before, directing most of the research programmes with central committees of "old guard", who, the younger researchers complain, are still just as likely as they ever were to fund work on political grounds.

Nevertheless, there is a revolution going on, albeit a rather quiet one. There are no tanks in the street and no protests outside the academy, but throughout Hungary's research community there are demands for the reforms along the lines of those that took place in the streets of Budapest itself 3 years ago. "It's a typical transition period for Hungary", says Lajos Nyiri, a former Hungarian science attaché: "It's absolutely clear that it is changing, but it is not clear where." It is also not clear when it will be resolved. One of the constants in Hungarian science is the ability of the academy to mire reforms — especially threatening reforms — in bureaucracy.

"Everything there runs at a palaeozoic rate of change", says Lawrence Cohen, environment, science and technology attaché for the US embassy. Outside the academy, Hungary has embraced Western ways. In the mid-1980s it had private businesses, commercial banks and even the makings of a stock market.

Even while it was a member of the Soviet bloc and looked east for trade, its historical ties to Austria (Vienna is just a few hours up the Danube) left it with one foot in the European door. Hungarian

industry also stayed close to international standards, although maintaining that productivity sometimes came at an economic cost to the country, which has the largest debt in Eastern Europe.

At the moment, however, Hungary is a troubled country. Its economy is a mess (inflation was 36 per cent last year), its people are, if anything, more dour than usual (it has the world's highest suicide rate; 4,235 Hungarians killed themselves in 1990), and its currency is still not stable enough to be convertible.

But it is also on its way up. Despite the collapse of the Soviet economy, Hunga-

HUNGARIAN SCIENCE AT A GLANCE

No. of full-time researchers	20,000
No. of academy research institutes	37
No. of universities	4
No. of specialized colleges	14

Total research/development funding	\$630m
Government funding*	\$320m

* Exchange rate 63 forint per dollar (1990 figures).

ry's traditional market, Hungary has the best trade balances in Eastern Europe and actually showed a trade surplus with the West last year. Goods — even Western goods — are plentiful and the only queues in Budapest are outside the Nike and Levis stores. Budapest is also home to Europe's most successful MacDonalds. Unemployment, despite the elimination of thousands of state-supported jobs, is still very low — less than 3 per cent. Wages are even lower, however, so many Hungarians have to do two jobs to keep their heads above water. The only place to get a live basil plant in Budapest is from a roadside stand operated by a agricultural engineer from one of the Academy of Sciences laboratories. She leaves her first job at about 4:00 every afternoon so she can get to her stand in time to meet the evening shopping rush. It is almost impossible to find a government office still functioning after 4:30.

Hungarian science is one of the bright spots in the country's future. It is the best in Eastern Europe (see graph on page 670) and it is strong in the applied sciences, an indirect result of declining government support of basic science over the past decade, which forced many of the university research institutions to find

support from industry. Although Hungarian industry cannot support much science in its current state, the experience of working in applied fields should improve Hungarian scientists' chances of finding work from Europe and the United States.

The language barrier, of course, remains a problem. According to academy statistics, only about a quarter of Hungarian researchers have passed competency exams in English, about the same number that passed Russian. After years of reviewing research proposals in-house, the Academy of Sciences and the government research fund (OTKA) are now starting to send them out for review by researchers in Austria and Finland to avoid longstanding charges of favouritism and bias in academy review. This means, of course, that the proposals must be written in English, something that may separate the old from the new guard. "The new generation and the ambitious speak English", says Géza Bittsánszky, head of the government's secretariat for science policy. "This may help trim the community."

Slimming the overpopulated and inefficient state-supported research machine is an oft-repeated goal, although there is little, short of underfunding, that the government is doing about it. Although everyone agrees that many of the academy research institutes were of poor quality, there has until recently been little in the way of impartial analysis to point out the culprits. A move to competitive grants and a real peer-review system, announced last year (see *Nature* 352, 745; 1991), will help. So will a parallel initiative to bring the universities to the Western model. The Athenaeum Initiative, one of the proposals now being considered, would allow universities to offer Western-style PhDs, a substantial departure from the current system in which the Hungarian Academy of Sciences decides who gets the advanced degrees that are necessary for tenured positions.

It is rather un-Hungarian for them to say so, but scientists admit that their current situation could in fact be a good deal worse. It could be Russia — indeed, the Soviet-style Academy of Sciences, with its ageing research institutes and academicians is for many researchers a reminder of just what was wrong with the bad old days. □

The more things change . . .

ON the corners of Nador and Zyiri streets in Budapest, there is an old yellow government building with a new sign. Inside, around 200 people administer the programmes of the Hungarian Academy of Sciences, just as they have for decades. But this is more than 2 years after the shift to democracy, and the government promise that such Soviet-style institutions would be radically reformed.

What has changed? The sign. What was once the "Secretariat" of the Academy is now simply the "Bureau". Otherwise the academy has the same president, secretary-general, 230 academicians (30



Old building, old ways.

of whom were installed by the Communist Party), and 200 administrators.

"It is the same people in different clothes", one staff member despairs. "If they cannot reform the government institutions, how can they reform the country?" She says that, as a patriot, she would be willing to lose her job for the sake of real changes, but that no one has even asked.

Academy officials say that it is unreasonable to expect faster change. "Naïve people do not understand how complicated these things really are", says György Odze, the academy's spokesman. "A reform should not always be radical." The problem in bringing in fresh blood, he says, is that it is difficult to avoid making change for its own sake. Many of the academicians are there because they were excellent scientists. Admittedly, many others won their position for more political reasons. But, Odze says, "it is not enough to tell 200 academicians: okay, boys, time to go home. We would have to pick new academicians. But how? It is very easy to cut off a man's head. It doesn't hurt and it only takes a moment. But then you have to worry about whether he was guilty."

One reform that everyone agrees on is the need for competitive, peer-reviewed research. That, at least, seems well on its way. Last year, for the first time, all Hungarian researchers were able to nominate peers to serve on new committees to review grant proposals for the Academy as well as OTKA.

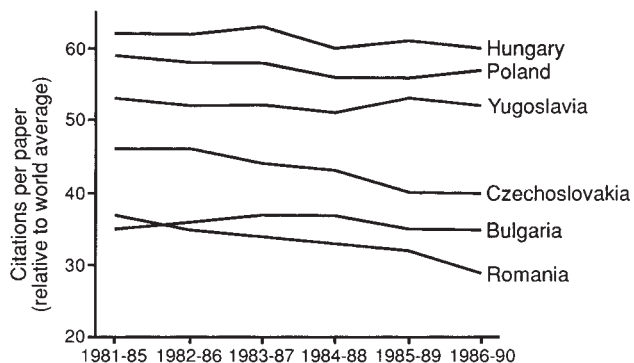
Where Hungarians do not all agree is on the ownership of the academy's 52

institutes, which employ about 6,000 scientists. The academy, naturally, wants to continue to control the facilities. There is nothing wrong with such academy-run institutions *per se*, Odze argues; the solution is to improve the way the Hungarian Academy exercises its control, not throw out the system altogether.

But many Hungarian researchers are not willing to wait. They want the institutes to transfer to an independent government agency, like the US or UK model of science agencies. Only then, they say, will the academy and its 'old boy network' of favour trading and personal loyalties cease to be the dominant force in Hungarian science. "It's absolutely clear that the academy has to become more 'academic'", says Lajos Nyiri, a former science attaché in the Hungarian embassy in Washington. He says that the Hungarian Academy has encouraged a "feudal" system of doing science in which researchers who gain the stature of "eminent scientists" decide who will be funded in the future. Although the recent moves to Western-style grants and universities (see *Nature* 352,

Hungarian researchers have consistently produced the highest quality research in Eastern Europe, although as a group they are still somewhat below the world average (100) and, like those of researchers in the other countries of the region, gradually declining relative to the international norm. In this analysis, citations to Eastern European research papers were counted in five-year 'windows'. Each window reflects the average of papers that were published and cited in that period.

(Source of figures: ISI science indicators database 1981-90.)



WESTERN INVESTMENT

Reality hits early hopes

Of all the emerging democracies, Hungary has had the most success in attracting Western companies and economic investment (more than \$1,000 million so far), thanks to its half-decade head start towards a free market and its strong base of intellectual talent. But a year ago, when the Gulf War broke out and the Hungarian government inflated the forint by 15 per cent in an attempt to make it convertible, the bubble of Western interest burst. Investment for the first 6 months of 1991 was less than half of the same period the year before. Most of the horde of US lawyers who invaded the country, hoping to broker sweet deals, have gone home again.

745; 1991) will help, getting the concept of merit-based research support to trickle down into the individual laboratories will take time.

Time, however, is not on Hungary's side. As with other Eastern European countries, a swelling 'brain drain' is a major concern. Every year, some 12 per cent of Hungarian scientists are working abroad; a quarter of those have been away for more than 5 years. The threat of losing a generation of young scientists seems to be the principal fear; in some fields the number of scientists in their 20s and 30s have fallen by 30 per cent.

Feudal laboratory management styles only make the problem worse. "The problem is of a promising young scientist with a stupid boss", says Nyiri. In the old Hungary, that boss might have friends in the academy to fund his programmes and secure his position. In the new Hungary, researchers will ostensibly have to compete on scientific merit alone. But one cannot legislate away 40 years of power structure. Even though there are now independent review committees in name, no one has forgotten the old loyalties. "These things take a long time to change," Nyiri sighs. "I am not optimistic that the young scientists will wait." □

The problem is mostly one of unreasonable expectations and the reality that change of Hungary's scale takes a long time. But the Western investors who went home empty-handed had legitimate concerns, too. A fully convertible currency is essential (it is now convertible for most business uses, although not yet for personal transactions) and the government must stabilize long enough to set the ground rules for foreign investment, especially in areas such as liability.

These things will happen eventually. And there are signs that the Western investors have not given up. Although no large deals have been signed since the Gulf War, a number of previous agreements —

NATURE · VOL 355 · 20 FEBRUARY 1992

including a \$250 million arrangement with General Electric to make light bulbs and almost \$400 million in separate investments by Suzuki, General Motors and Ford to make cars in Hungary — are still going through. PlanEcon Research Associates, a Washington-based consulting company, estimates that Western investment in the Eastern European nations will rise by half this year, to about \$7,800 million.

But for the most part, the investors are looking at Hungary as a cheap labour pool or as a market. One of the few deals that would have invested in Hungary's high-technology industry — a plan for Groupe Bull, the French electronics company, jointly to develop and produce consumer electronics in Hungary with Videoton, a large Hungarian telecommunications company — went under when Videoton declared bankruptcy late last year, putting more than 1,000 engineers and tech-

nicians out on the street.

Part of the problem is simple economics: investments in Hungary will generally take years to bear fruit and, as the case of Videoton demonstrates, can be risky. But such chances are not uncommon. What makes Hungary especially difficult for investors now is the still uncertain laws on such basic business issues as environmental responsibility and ownership. Few companies would buy a Hungarian business if they knew that they could be sued for chemical contamination that happened decades ago under a Communist regime. And what happens if someone comes forward and claims that land purchased by a Western investor had belonged to their grandparents before it was nationalized in the 1940s? The law is not yet clear on any of these points. And parliament, with an economic revolution on its hands, is not likely to resolve these issues soon. □

TECHNOLOGY TRANSFER

Digging out of the ruins

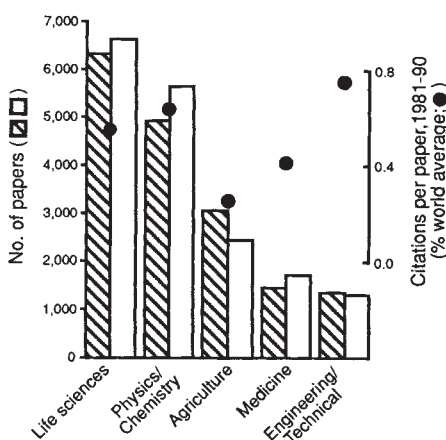
FOR 40 years, being a Hungarian researcher meant a safe job, steady if unspectacular funding, and a captive market for any products one might come up with. That, of course, has all changed, nowhere as much as in the transfer of science to commerce. Where there was once a market within the Soviet bloc for research that was even a decade behind the West, inventions must now compete directly with the rest of the world's.

Despite the reputation of a state owned research enterprise, Hungarian researchers have been doing more over the past four decades than simply holding down their jobs. After all, Hungary's science is the strongest in the region, and applied and technical research is what the country does best, even if not quite as well as the world as a whole (see chart). So last year the government set up an office — known as the Office for International Technology Cooperation (NTEI) — to try to sell, license or find partners for as much domestic science and technology as it possibly can.

Last April, NTEI published a booklet entitled *Hungarian Ideas Awaiting Foreign Partners* that shows just how tough the office's job will be. Of the booklet's 22 technologies, most are agricultural — freeze-resistant yeast, genetic engineering techniques for fish and the like. Hungary has always been something of a breadbasket for the region, and its agricultural innovation may well be excellent. But outside agriculture, the cream of Hungarian technology (at least as offered for sale) gets a bit thin. "Medical check-up of the population" is actually a plea for help in putting together 10 million cheap tests for cholesterol and blood sugar. And the "method for loss-

reduction of transmission lines" is really just an advisory committee with a big problem: Hungary needs more power, but anti-nuclear activists will not let it build more nuclear reactors and environmentalists will not let it complete a joint Hungary-Czechoslovakia hydroelectric dam on the Danube. Short of building a conventional power plant that it cannot afford to run, the only solution is to cut electrical losses in the power grid.

This seems to be a common problem with Hungarian technology and technical



Life-science research is what Hungary does most of, but engineering, technology and applied science is what it does best, according to this study. Although the applied disciplines are a small part of the research enterprise, they appear still to be reaping the rewards of Hungary's traditional strength in physics and chemistry (which, while no longer up to international standards, are still one of the better areas of Hungarian science). Hatched boxes, 1981-85; open boxes, 1985-90.

(Source of figures: ISI science indicators database 1981-90.)

opportunities: after 6 months of operation, the new office had yet to make a single sale. Selling businesses is even harder. The State Property Agency has 2,180 state-owned company businesses that it must privatize: by late last year, only two companies had been sold, although others have attracted some foreign investment.

Zsolt Köhalmi, director of NTEI, harbours no illusions about the challenge before him. The ideas, he says, "are not the greatest miracles on Earth, but they're not bad." From a modest two-office suite on the third floor of an old government office building, Köhalmi is assembling a database of such offerings. He now has a list of more than 1,000. The problem, he says, is not a lack of Hungarian innovation, but unrealistic expectations from Western investors. "Most of these people discovered Eastern Europe when the Berlin wall came down in 1989", he says. They generally come on the expectation that there are some good ideas in Hungary that need developing because of Communist inefficiency, which Köhalmi says is largely true.

"But how fast can you dig something out of the ruins of Communism?" he asks. Because Soviet bloc policies ensured that Hungarian scientists could find a market for almost anything they produced, there was little feedback from customers. Good ideas were often buried in half-baked products. Western-style market forces were virtually nonexistent.

So if a Western company does find a Hungarian technology with promise, it should not be surprised to find it a little more rough around the edges than the sort of technology offered by a typical US university. It may have to spend more time developing it into a product, which is inconvenient, Köhalmi concedes. But his prices are low, and he promises that the benefits of a cooperative relationship with Hungarian scientists will pay off eventually.

After an initial surge of interest, however, many prospective buyers seem to be unwilling to cope with the hassle of a cooperative agreement. When US companies do find a Hungarian idea that they think is worth investing in, Köhalmi complains, "they often decide it is cheaper to just use the good old brain drain — to bring the scientists to the States."

But those investors who do make the extra effort may find payoffs where they least expected them. After all, it was a Hungarian invention that put out many of the oil-well fires after the Gulf War. It may not have been the most high-technology device — it consisted, in fact, of a pair of jet engines strapped to a tank — but it was quite effective and, somehow, typically Hungarian. □

West's gift becomes a model

IN a year in which 6,000 visitors passed through its doors, a new centre to bring Western environmental research and tactics to Eastern Europe is quickly becoming the focus of the region's budding environmental programmes — and their critics. Called the Regional Environmental Center (REC), the Budapest-based centre has drawn fire for, at one time or another, being too green, not green enough, too American, too Hungarian, too high-tech, and not high-tech enough. This is, says Lawrence Cohen, science and environment officer in the US embassy in Budapest, very reassuring; if everyone thinks the REC is not responsive enough to their needs, it shows, at least, that no one faction has hijacked the centre.

By most accounts the REC is thriving, as a horde of scientists, activists, government officials and companies turn to it to set up environmental surveys, grassroots non-governmental organizations, new environmental legislation and remediation campaigns.

The REC came about when the US embassy in Budapest was trying to think of something to give Hungary on the occasion of President Bush's visit in 1989. First planned as an effort to help establish environmental programmes in Hungary, its reach was extended to all of Eastern Europe as the region became democratic. Although it started with the United States alone, it is now equally supported by the European Communities (combined US and EC funding is \$10 million for 3 years) and, to a lesser extent, by Canada, Hungary, Austria, Japan, Finland, Norway and the Netherlands. Total funding is about \$20 million for 3 years.

In its first full year of operations, the centre has been deluged. It has become an important stop for visiting Western environmental researchers, as well as playing host to policy-makers, industry representatives, members of environmental and other pressure groups and students.

Most of the centre's work is now collecting and disseminating information — a quantity that has traditionally been in short supply in Eastern Europe. It is setting up a database of environmental data for the region, as well as a computer network to link the Budapest centre with small offices in several other cities in the region, where local researchers can find information and ask questions in their own language. The REC is also developing two mobile laboratories which will serve as working laboratories and classrooms to demonstrate monitoring and computer modelling equipment.

In one typical encounter last year, the REC was approached by members of

environmental groups who discovered a huge mercury contamination in eastern Hungary and asked for help. The centre invited a group of US researchers to visit and give technical advice. While the specialists were in Budapest, they prepared a report to parliament to help in developing future legislation. This year the US Peace Corps will assign 36 environmental scientists and 20 computer scientists to work through the centre in the region.

Although pollution problems in Eastern Europe are among the worst in the world, centre officials say that the state of environmental politics may be just as serious. This is partly because few in the region understand the scope of the contamination. "The realization of how great



On the Danube — environmental confusion in Hungary's government.

the region's environmental problems are may be greater in the West than in central Europe itself", says Branko Bosnjakovic, REC's outreach and development manager. "People here still think that they cannot afford to protest, for fear of losing their jobs."

With the prospect of better economic times, people who have lived in Eastern Europe all their lives are reluctant to slow growth with environmental measures, or to make personal sacrifices. "Now that it is possible for people to buy cars as a status symbol, they naturally want to, which means less incentive to clean up the public transportation system, and more polluting cars on the road", says Bosnjakovic.

Governments, in turn, are afraid to impose strict environmental laws, lest they discourage Western investment. And now that each nation sets its own policies, rather than taking its cues from Moscow, agreeing on environmental legislation is nearly impossible, even though most of the region's legislatures have asked the REC for advice.

Western companies contemplating investing in the region also consult the REC. They are most concerned about existing contamination, with the possibility that they might be unsuspectingly

taking over an undocumented toxic waste site the principal nightmare.

Unfortunately, the REC's attempts to set up cooperative and environmental research teams have been complicated by national and ethnic rivalries. "I continue to try to figure out what 'neutral' means here. What line do I toe?", wonders Stephen Wassersug, REC programme manager. At one meeting last year, representatives from non-government organizations in the two of the Yugoslavian republics refused to speak to each other. Antagonism extends to environmental officials in the new governments as well. Eastern European activists like to say that many environmental officials in their region suffer from what they derisively call "watermelon syndrome" — green on the outside but red on the inside.

Ironically, many of the political movements that brought about the rise of

democracy in the region started in environmental organizations, which were among the few activist groups permitted under Communist rule. Once it became clear that the old regimes were going to fall, many of the dissidents shed their environmental cloak and campaigned openly as political candidates. This was good for the rise of democracy, of course, but it took its toll of the infant environmental movement. "After the change, people were concerned with the massive economic problems", says Bosnjakovic. "Environmental issues slipped on the political agenda." Now, 3 years later, that movement is rebuilding with people motivated by environmental rather than political concerns.

Despite the political muddle, Hungary has announced its intention to reach EC levels of environmental legislation by the mid-1990s. This is "wildly optimistic", says Cohen, "They're trying to build an entirely new environmental code at the same time that they're dealing with privatization and all the rest." The US Environmental Protection Agency has lent Hungary experts to help legislation, but most of them consider Hungary a decade away from draft EC compliance. They agree, however, that without the REC it would be even longer. □

Unrequited search for heavy neutrino

Is there some impending crisis in the relationship between particle physics, neatly but incompletely described, and the qualitative properties of the Universe at large?

THE general belief that most of the mass of the Universe, perhaps four-fifths of it, consists of matter than cannot be seen (and which is thus called 'dark') is apparently undimmed by the equally general failure to tell what it may consist of. No sooner, it seems, does a candidate constituent make its appearance in the literature than there emerges a group of Jeremiahs eager to demonstrate that the candidate will not measure up to the role that astrophysics and cosmology require it to play. Not much has recently been heard of the axion as a generalized cement that will hold the Universe together. Is the idea that a neutrino with a mass equivalent to about 17,000 electron-volts (17 keV) also now destined for oblivion?

It is a curious tale. The idea of the 17-keV neutrino goes back to J. J. Simpson in 1985, which is why it is sometimes called Simpson's neutrino. The original evidence was that of a kink in the energy spectrum of electrons from the β -decay of tritium at an energy of about 17 keV, interpreted as the emission of neutrinos with such a mass in partnership with the electrons whose energy had been measured. Simpson estimated that heavy neutrinos might account for about 3 per cent of the total. More accurately, while most decays of tritium nuclei would yield an electron and an ordinary neutrino, 3 in 100 would yield an electron and a heavy neutrino.

Since then, several groups have sought to replicate the findings by studying the β -decay of other nuclei. In a particularly elegant measurement based on the incorporation of ^{14}C into a sample of germanium wired to function as an electron detector, Sur *et al.* from the Lawrence Berkeley Laboratory produced evidence to support Simpson's conclusion, suggesting a 1 per cent admixture of heavy neutrinos in the decay products of ^{14}C (*Phys. Rev. Lett.* **66**, 2, 444; 1991). Similar conclusions were reached at Oxford from measurements of ^{32}S (Hime, A. & Jelly, N.A., *Phys. Lett.* **B257**, 441; 1991).

Unfortunately, such evidence is hardly ever compelling. People looking for a small kink in a measured curve necessarily contaminated by background must wave their hands vigorously if they are to carry conviction. Direct detection of the particle necessarily carries greater weight. The trouble, here, is that the particle is a neutrino. So what, if it exists, is this neutrino?

Like the others, it is electrically neutral and capable of interacting with other forms of matter only weakly. As particles go, the mass is tiny — a few per cent of the electron mass — but that is far greater than the mass required to close the Universe if it were supposed that all extant neutrinos have the same mass. Simpson's neutrino, if it exists, must be a rarity among neutrinos. The manner of its first discovery also suggests that it belongs to one of the families of neutrinos linked to other particles (and, indirectly, to each other) by the weak nuclear interaction. That seems generally agreed. But not much else.

One obvious difficulty is that Simpson's neutrino cannot be either of the two particles associated with the electron and the muon in the standard scheme of things; the mass is too great for that. But, for the past two years, people have been reasonably confident that there are only three families of two neutrinos (an ordinary and an anti-neutrino), linked with and characteristic of one of three lepton-pairs — electrons, muons and taus. If Simpson's neutrino is neither the electron nor the muon neutrino, it must be that associated with the tauon.

But that raises problems. James M. Cline and Terry P. Walker from the Ohio State University at Columbus now argue that, if Simpson's neutrino is the tauon-neutrino, there are difficulties about its lifetime (*Phys. Rev. Lett.* **68**, 270; 1992). On the standard model of weak nuclear interactions, the neutrino should be stable against radiative decay, with a lifetime of about 10^{22} seconds. But, the argument goes, if the microwave background radiation is not to be "unacceptably" distorted, the lifetime of Simpson's neutrino must be less than 10^6 seconds (a couple of weeks or thereabouts). The gap between the two limits seems unbridgeable.

Cline and Walker, with admirable cheerfulness, chase the issue into a cul-de-sac by supposing that Simpson's particle does not have the properties described by the standard model of the weak interaction, but that it may be something different. But even with that extra freedom, it seems that Simpson's neutrino cannot be easily accommodated. Several cosmological considerations constrain the properties of the 17-keV neutrinos. For example, reconstructions of the formation of ^4He in the early Uni-

verse are possible only if the Simpson particle is accompanied by another electrically-neutral weakly interacting particle, but apparently inescapable difficulties still arise in reconciling expectation with reality. Further arguments with similar conclusions attend the discussion of the emission of neutrinos from the five-year supernova SN1987A.

What Cline and Walker modestly conclude is that, "if Simpson's neutrino survives, it surely will point to exciting and surprising new particle physics in the neutrino sector". Their readers will be forgiven for asking whether that state of affairs may not already have been reached, even if nobody has yet been able to define what the "new particle physics" consists of. The simple way out, of course, would be to insist that the laboratory experiments are faulty, or that their admittedly hazardous interpretation is wrong, but that would be unpardonably unfashionable and dangerous as well.

Yet the argument continues. Enqvist, Kainulainen and Thomson (from Copenhagen, Copenhagen and Manchester respectively) argue that there would be advantages if the lifetime of the Simpson neutrino were greater than the decoupling time of electron neutrinos in the supposed Big Bang, but less than the time at which nucleosynthesis began (*Phys. Rev. Lett.* **68**, 744; 1992). Gorski, Silk and Vittorio (from Princeton, Berkeley and Aquila in Italy respectively) argue more broadly that the issue of whether the doctrine of cold dark matter can be reconciled with what is known of the anisotropy of the microwave background will be settled (or otherwise sharpened) by the latest analysis of data from the Japanese COBE satellite, due out any day (*Phys. Rev. Lett.* **68**, 733; 1992), but that little can be done until then to reconcile the Simpson experiments with their cosmological implications.

Where all this will lead is anybody's guess. It will be a remarkable development if a sketchy laboratory observation of irregularity in an experimental curve leads directly to an upheaval in cosmology, but this year is, after all, the fourth centenary of the occasion when Galileo first set his heliocentric cat among the Vatican's pigeons. It is possible, of course, that Simpson's neutrino will not "survive" the year, but otherwise it is in everybody's interest to decide what the dark matter is made of. **John Maddox**

Alpha-particle after effects

H. John Evans

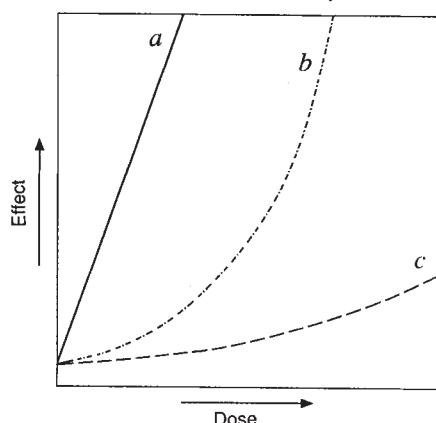
ON page 738 of this issue¹, M. A. Kadhim and colleagues describe experiments which, unexpectedly, show that α -particle irradiation of a cell may induce a transmitted genetic instability that may well result in the generation of genetic abnormalities later in the same cell line. Given that the genetic effects of low levels of ionizing radiation remain of considerable concern in terms of environmental health, the findings are likely to provoke something of a stir.

There has long been debate over the relative biological efficiencies (RBE) of ionizing radiations of differing qualities in inducing genetic damage at equivalent low exposure levels. With radiations of high linear energy transfer (LET), the ionizations in the irradiated cell are clustered as dense tracks; this pattern of energy deposition is characteristic of cosmic rays, the 'knock-on' protons from fast neutrons or the α -rays from plutonium from nuclear installations, and the decay products of the radon gas that form a major part of the natural background radiation². Radiations of low LET on the other hand, such as X-rays or the γ -rays from the atmosphere, soil and buildings, not to mention food, release their energies as sparsely distributed small clusters within the cell. These differences in spatial energy transfer are reflected in the different propensities of these radiations to induce genetic changes resulting in inherited defects or carcinogenesis.

Early studies, such as those using insect eggs or mammalian cells in culture, in which it was possible by selective lead shielding to irradiate specifically the cell nucleus or the cytoplasm, showed that the passage of a single α -particle track through a nucleus was sufficient to kill a cell. However, the protons produced by fast neutrons or the α -particles from plutonium may of course deposit their energy in the near vicinity, or at the periphery, of a nucleus, and so may not always be lethal but may still result in genetic damage. Indeed, much is known of the RBEs of radiations of different mean LETs in inducing cell death, chromosome aberrations or single gene mutations in human peripheral blood lymphocytes or fibroblasts irradiated *in vitro*.

In very broad terms, with low LET radiations the frequency of induced genetic damage is related to dose and dose rate. At high rates of exposure, chromosome damage increases with increasing dose roughly in proportion to the square of the dose, giving a curvilinear dose/effect response. If the radiation

is given at low dose rate, the response curve may approximate to linearity between dose and effect, so that damage induced by low LET radiations may be repaired and their repair is time dependent. In contrast, with very high LET radiations little repair seems to be possible, the aberration frequency increases linearly with dose and the effects are not reduced by decreasing dose rate (that is, increasing exposure time). It follows that, because of differences in the shapes of the response curves, high LET radiations are relatively far more



Dose-response curves for the induction of chromosome damage in human lymphocytes. *a*, Fast neutron, for both acute and chronic exposure; *b*, X-rays, for acute exposure; *c*, X-rays, for chronic exposure.

effective than low LET radiations at low levels of exposure, and that the RBE decreases with increasing levels of damage. (See figure.)

Kadhim *et al.*¹ now report that not only are there quantitative differences between the efficiencies of X-rays and α -particles in inducing genetic damage, but that there may also be hitherto unsuspected qualitative differences. The authors employed an *in vitro* assay for the survival of colony-forming stem cells in the bone marrow of mice, after exposure of these cells to a range of doses of α -particles (0.25, 0.5 and 1.0 gray) from plutonium-238 or to a single X-ray dose (3.0 gray). Alpha particles have very poor powers of penetration, so the cells were exposed in thin-layer suspensions and then plated into Petri dishes; survival was determined by the ability of surviving cells to give rise to colonies and chromosome analysis was performed on 7-day-old colonies.

At survival levels of about 40 per cent or 20 per cent, the mean number of α -particles traversing a cell of 7 μ m diameter was respectively 1 or 2. Many surviving cells would therefore not have been traversed by a particle and others

may have suffered ionizations from only a small part of a track. When the chromosomes of these surviving cells were analysed a surprisingly high frequency of chromatid-type aberrations was evident; these aberrations were non-clonal and were evidently engendered many cell divisions after the original exposure. In contrast, the 5 per cent of cells surviving an exposure to 3 gray of X-rays showed a few clonal abnormalities, presumably engendered during the first cell cycle following exposure, but nonclonal chromatid-type aberrations were present at the expected control levels.

These findings imply that, at not too dissimilar survival levels, α -particle exposure may induce damage that is transmitted to daughter cells and which may result in the expression of a genetic instability in later cell generations. The generation of new chromatid aberrations many cell cycles after exposure is known to occur after exposure to certain chemical mutagens and, indeed, there are a number of conditions in humans where an innate chromosome instability is a consequence of the inheritance of a single mutant gene³. It is well known that such individuals are more prone to the development of early cancers, and Kadhim *et al.* point to the possible importance of their findings in radiation-induced leukaemogenesis and childhood leukaemia clusters associated with nuclear sites. Notwithstanding the fact that clusters of childhood leukaemias occur at locations other than those associated with nuclear installations⁴⁻⁶, how important are these new findings in the general context of radiation exposure?

One concern for radiation protection is the nonuniformity of absorbed dose. For instance, it is difficult to evaluate the consequences of the deposition of a 'hot particle' in a sensitive tissue. An α -emitting speck of plutonium will have a high probability of killing cells in its immediate vicinity, but if it engendered instability in nearby surrounding surviving cells its effect would be very much greater than that predicted on the total body dose received. A connected concern is the RBE of high versus low LET

1. Kadhim, M. A. *et al.* *Nature* **355**, 738-740 (1992).
2. Clarke, R. H. & Southwood, T. R. E. *Nature* **338**, 197-198 (1989).
3. German, J. (ed.) *Chromosomes and Cancer* (Wiley, New York, 1974).
4. Evans, H. J. *BioEssays* **12**, 541-549 (1990).
5. Evans, H. J. *Nature* **347**, 712-713 (1990).
6. Openshaw, S. & Craft, A. in *Studies on Medical and Population Subjects*, 53: *The Geographical Epidemiology of Childhood Leukaemia and non-Hodgkin Lymphomas in Great Britain 1966-83* (ed. Draper, G.) 109-120 (HMSO, London, 1991).
7. Committee on the Biological Effects of Ionizing Radiation BEIR IV (National Academy of Sciences/National Research Council, Washington, DC, 1988).
8. Henshaw, D. L., Eatough, J. P. & Richardson, R. B. *Lancet* **335**, 1008-1012 (1990).
9. Radiation Effects Research Foundation Update 3, Issue 3 3-6 (US-Japan Radiation Effects Research Foundation, Hiroshima, 1991).

radiations at low levels of exposure. The α -particles from radon and its daughter products are known to interact with the effects of cigarette smoke in inducing lung cancers in uranium miners⁷, and a possible association between leukaemia prevalence and background levels of radon has been suggested by Henshaw and his colleagues⁸.

It is relevant, however, that between 1 and 2 per cent of the radiation dose received by the survivors at Hiroshima consisted of fast neutrons, and recent simulation studies show that the RBE of these high LET (0.21 MeV) radiations relative to cobalt-60 γ -rays is somewhat higher than previously estimated, and may be as high as 60:1 at low dose levels⁹. In other words, neutron-induced damage at low doses was at least as great as damage caused by sparsely ionizing

tracks. Because data from the atomic-bomb survivors provide a substantial input into estimates of radiation hazard, the possible consequences of induced genetic instability will, unknowingly, be included in these estimates, and the effective biological doses of the high LET radiations will be somewhat larger than those previously assumed. The difficulties of extrapolating from effects at low LET to those at high LET is further underlined by the studies of Kadhim *et al.* on mouse bone-marrow cells — studies which, if confirmed, raise questions both as to the mechanisms involved and of the relative importance of these delayed effects. □

H. John Evans is in the MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK.

DYNAMICAL SYSTEMS

Coming to grips with chaos

Francis C. Moon

CHAOS, the study of unpredictable dynamics in nonlinear systems, has focused attention on understanding the transition from regular to complex chaotic motions since the topic first developed over a decade ago. Recently, chaos researchers have looked at how to tame chaotic behaviour using the very nonlinear properties that lead to unpredictability. Examples of how 'nonlinear thinking' can be used to invent controllers for systems with complex dynamics were the highlight of a meeting* on experimental studies of chaos.

The hallmark of chaos is said to be the sensitive dependence of the evolution of a system on the precise details of its initial conditions: slight changes in the physical circumstances can lead to wildly differing outcomes. A classic example is the erratic oscillation of a real pendulum driven by a sinusoidal oscillator. Unlike a simple oscillatory system, which will repeatedly and predictably return to its original state, a chaotic system will sample a wide range of possible states, repeating none, defined by a 'strange attractor' on the dynamical map used to describe evolving systems.

The most interesting experiments described at the meeting have built on the theoretical ideas espoused by E. Ott, C. Grebogi and J. A. Yorke in their paper *Controlling Chaos* (*Phys. Rev. Lett.* **64**, 1196–1199; 1990). Ironically, they proposed exploiting the very stochastic (random) characteristics of a chaotic system to lock it onto periodic behaviour through the unstable periodic points embedded in the strange attractor (the

points mathematically represent periodic behaviour but are dynamically unstable).

M. L. Spano and W. L. Ditto (Naval Surface Warfare Center, Maryland) described how the use of small perturbations enables one to switch between different orbits embedded in a strange attractor using a control scheme derived from the underlying map for the attractor. Their experiment concerns the vibrations of a flexible beam made of an amorphous ferroelastic material called Metglas ($\text{Fe}_{81}\text{B}_{13.5}\text{Si}_{3.5}\text{C}_2$), whose effective Young's modulus is sensitive to small magnetic fields. When an applied field is periodically perturbed, the beam vibrates chaotically. Embedded in the attractor is an infinite set of unstable periodic orbits. The position of the beam was monitored optically, and the results fed into a computer which, initially, calculated a nonlinear map for the beam's chaotic attractor. Then it was possible for the computer to sense when the beam oscillations were within a prescribed amount from a fixed point in the attractor, and to adjust the mean applied field for a single cycle. In this way, Spano and Ditto could stabilize an otherwise unstable periodic orbit, and switch between different fixed points (Fig. 1).

E. R. Hunt (Ohio University) does the same in an electronic circuit system consisting of a rectifier-type diode in series with an inductor driven by a sinusoidal voltage (Fig. 2). Again, the circuit oscillates chaotically, this being

apparent from the peak output current on each cycle of the drive frequency. The resulting map generated by the strange attractor is then used to produce a control signal. Using this method, Hunt could stabilize almost all uncontrolled unstable orbits up to period 23. These two experiments employ essentially digital control schemes. But H. H. Bau and J. Singer (University of Pennsylvania) use an analog control method to quench Lorenz-like convection chaos in a ring-like tube of fluid heated from below, called a thermosyphon. Without the control, the fluid rotates in the tube, switching from clockwise to anticlockwise motions in an unpredictable manner. Using temperature measurements as a feedback signal, their controller can produce steady or periodic motions. It can also be used as a switch to stabilize and change from clockwise to anticlockwise fluid rotations.

An unusual extension of this idea of controlled chaos was to link together nonlinear oscillators to achieve chaotic synchronization, perhaps as a means to secure communication systems against eavesdropping. L. M. Pecora and T. Carroll (Naval Research Laboratory, Washington, DC) described how one might 'paste together' nonlinear subsystems to build useful chaotically driven

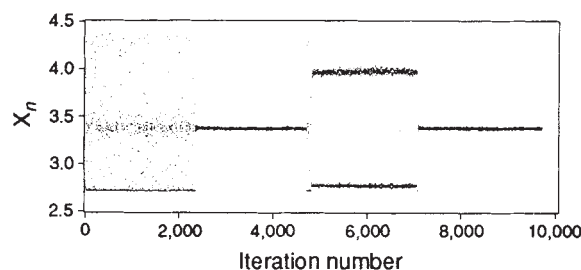


FIG. 1 Time-series data from Spano and Ditto's oscillating magneto-elastic beam. At first, the deflection X varies chaotically, but the period-1 and -2 orbits can be stabilized once the chaotic attractor is known. (From W. L. Ditto, S. N. Raueo & M. L. Spano *Phys. Rev. Lett.* **65**, 3211–3214; 1990.)

systems. In one example they built a pair of periodically driven nonlinear circuits whose periodic outputs were out of phase. By adding signals to each circuit from a third chaotic circuit with a strong periodic peak and a weak chaotic spectrum, they were able to bring the two original circuits into phase synchronization. An experimental realization of this idea was also described by J. Gullicksen (Loral Aerospace) and colleagues (University of California, Berkeley) who generated a chaotic carrier in a subsystem of two coupled digital phase-locked loops. A single loop (the receiver) was then synchronized with the chaotic output of the two-loop subsystem (the transmitter). The next step will be to introduce modulation on the chaotic carrier which will securely transmit information. Besides communication applications, Pecora and Carroll suggested possible

* First Experimental Chaos Conference, Arlington, Virginia, 1–3 October 1991. (Proceedings to be published by World Scientific.)

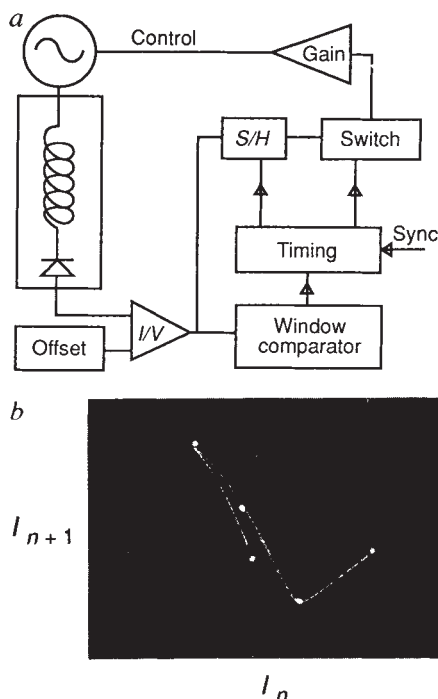


FIG. 2 *a*, The circuit used by Hunt to demonstrate control of chaos. A sinusoidal oscillator (top left) drives a coupled diode and inductor (box), whose output is monitored by a sample/hold device (S/H) triggered by a comparator. When the peak current falls within a preset window, it is used to modify the signal generator (I/V, current-to-voltage converter). *b*, A first-return map for the circuit, showing how the current on each drive cycle corresponds to the current on the previous cycle. Without control, the map shows the chaotic attractor. With control, five overexposed points appear; these points correspond to a stabilized period-5 orbit. (From E. R. Hunt *Phys. Rev. Lett.* **67**, 1953–1955; 1991.)

applications in the areas of robotics and bioelectronics.

There are probably two aspects to the oxymoron 'controlling chaos'. In the simpler case one seeks to adjust some control parameter either to quench a chaotically dynamic system or to produce chaotic dynamics in an otherwise quiet or periodically behaved system. But the new concepts inspired by Ott *et al.* and illustrated at the conference show how one might constructively use the exponentially divergent nature of chaotic orbits and the extreme sensitivity of these systems to small perturbations to design useful systems. These recognize that systems with nonlinear responses exhibit numerous basins of dynamic behaviour, which can be exploited by clever engineers and applied scientists to produce more creative solutions than those that are based on linear dynamic systems. □

Francis C. Moon is in the Sibley School of Mechanical and Aerospace Research, Cornell University, Ithaca, New York 14853, USA.

PALAEOCLIMATE

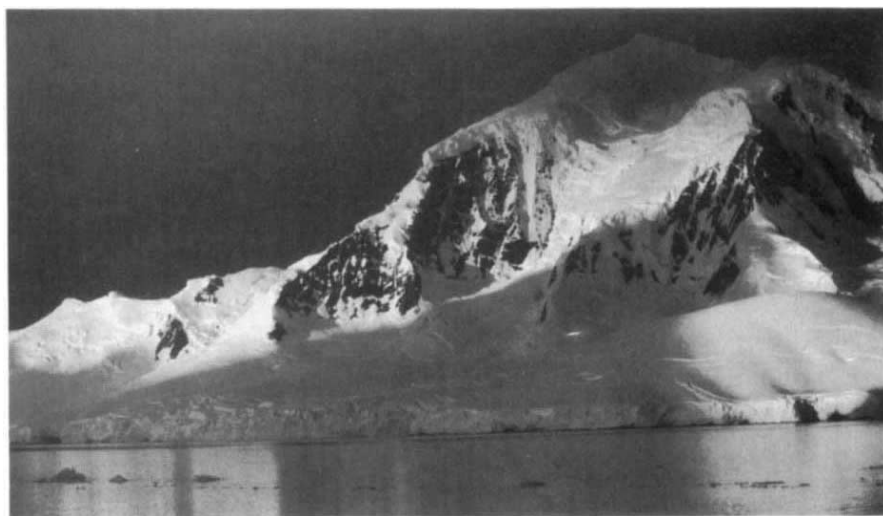
Paradox lost and paradox found

James F. Kasting

IN the stepwise pursuit of the truth, scientific theories evolve and sometimes change direction. Rarely, however, do they come full circle in the way that palaeoclimate theory has done in the past few years. Twenty years ago, Sagan and Mullen¹ pointed out an interesting puzzle in understanding the Earth's climate history: theoretical models suggest that solar luminosity was substantially lower in the past; consequently, the Earth should have been globally glaciated more than two billion (10^9) years

positioned adjacent to the South Pole and by reinvoking the idea of low solar luminosity. So, reduced solar luminosity in the past, once perceived as a problem, has been recast as a solution; meanwhile, increased atmospheric CO_2 in the past, once seen as a solution, is now viewed as a problem.

Whether or not high CO_2 is a problem for the late Ordovician, of course, depends on how strongly one believes in the predictions of Berner's model of the geochemical cycle⁵. Those who have ex-



Ice sheets might grow despite greenhouse CO_2 , if continental land masses abut onto the poles. (Picture reproduced from *Wild Ice*, Smithsonian Institution Press; 1991.)

ago had other factors remained constant. The geological evidence, however, suggests that this never occurred. This 'faint young Sun' paradox, as it came to be called, was resolved in the minds of at least some climatologists by the idea that the amount of atmospheric CO_2 was significantly higher than today^{2,3}.

Now, Crowley and Baum⁴ pose a new climatic paradox in connection with the late Ordovician glaciation, which occurred on the southernmost part of the Gondwanaland supercontinent 440 million years ago. Theoretical models⁵ predict that atmospheric CO_2 concentrations were high (about 13 times present levels) at that time, as a consequence of increased volcanism and decreased rates of continental weathering. Weathering sequesters CO_2 in carbonate sediments; volcanism recycles these sediments, restoring CO_2 to the atmosphere. Consequently, the Earth might be expected to have been too hot to have ice caps. Yet, the area of Gondwanaland covered by glaciers is thought to have been comparable to that of present-day east Antarctica. Crowley and Baum resolve this new paradox by postulating that the supercontinent was kept cool by being

permented with similar types of models realize that a great many uncertainties are involved in calculating past CO_2 levels. The long-term source for CO_2 , metamorphic degassing of sediments containing carbonate minerals or organic carbon, is assumed in Berner's model to be linearly related to seafloor spreading rate, which itself is inferred from inversion of the sea-level record⁶.

The accuracy of this approximation, and of the sea-level curve itself, is difficult to evaluate. The long-term sink for CO_2 , weathering of silicate rocks on land followed by deposition of carbonate and organic carbon in marine sediments, is governed by rate laws that incorporate the effects of globally averaged temperature and geography on surface runoff, as well as the effects of temperature and land plants on chemical weathering rates. Land plants are thought to reduce atmospheric CO_2 levels by pumping CO_2 down into the soil, thereby enhancing the rate of chemical weathering. This effect is parameterized in Berner's model by a formulation, developed by Volk⁷, that attempts to account for the effect of CO_2 on plant growth and metabolism. Because vascular land plants did

not evolve until the Silurian, which followed the Ordovician, this relatively complex weathering rate law is replaced by one that has a linear dependence on atmospheric CO_2 . Aside from being arbitrary, this approximation ignores the possible presence of widespread microbial mat cover, a phenomenon that Schwartzman and Volk have suggested⁸ might be very effective in drawing down atmospheric CO_2 .

Berner's own estimate for the uncertainty in predicted CO_2 concentrations during the late Ordovician is at least a factor of 5, from 4 times to over 20 times the present level. If the lower bound of his CO_2 estimates is used, along with an estimated 4 per cent reduction in solar luminosity⁴, the net radiative forcing of the atmosphere would be approximately the same as today⁹, and the paradox of the high-latitude glaciation effectively disappears. Conversely, if the upper bound on CO_2 pressure is used, it is virtually impossible to produce continental glaciation using Crowley and Baum's model. Indeed, the latter problem would be further exacerbated had Berner included low solar luminosity in his own calculation. This would have lowered surface temperatures and reduced weathering rates, allowing atmospheric CO_2 to climb even higher than Berner calculated.

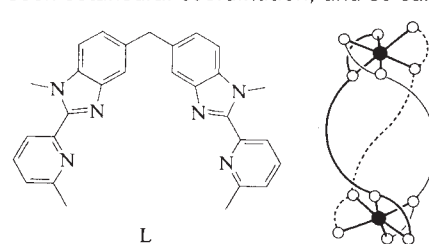
The other half of Crowley and Baum's glacial climate theory involves the importance of continental positioning and seasonality on ice-sheet development. Their idea, which was originally developed to explain the ice-free mid-Cretaceous¹⁰ (around 100 million years ago) is that ice sheets are difficult to form when a large continent is positioned directly over the South Pole, because the warm summertime temperatures prevent the interannual accumulation of snow and ice. The growth of ice sheets is easiest when the continent is positioned tangential to the pole (so that the pole lies on a coastline) because the high thermal inertia of the cold polar ocean keeps the summers relatively cool. This concept is demonstrably effective in Crowley and Baum's energy-balance climate model and has a certain degree of merit on purely logical grounds. Whether the late Ordovician glaciation provides a critical test for this idea is debatable, though, as one might explain the observation using CO_2 and solar luminosity alone.

What the Crowley and Baum paper really points out is the need for a CO_2 palaeobarometer that extends back beyond the past two hundred thousand years. We can never hope to understand more ancient climates without some reliable means of estimating past atmospheric CO_2 concentrations. Several groups, some of them mentioned by

Triple helix in a trice

NATURE loves nothing more than a helix, and what can be done once can be tried three times over. In *Angewandte Chemie (Int. Edn Engl.* **30**, 1490–1492; 1991) Alan F. Williams and colleagues at the University of Geneva describe the synthesis of a triple-helical complex held together by a pair of cobalt ions. The synthesis of helically coordinated organometallic complexes was pioneered in the laboratory of Jean-Marie Lehn (see *Nature* **346**, 339–342, 314–315; 1991) who uses copper ions to lock together pairs of coiling ligands. The same philosophy is now employed by Williams and colleagues. Each ligand, a sinuous chain of aromatic rings and heterocycles (see below), has four binding sites (nitrogen atoms)

but is too inflexible to wrap entirely around a single metal centre. The cobalt ions seek octahedral coordination, and so can bind to two of the four coordination sites



on each of three ligands. The free ends of the ligands can then latch onto a second cobalt ion. The authors found that mixing a solution of the ligand with cobalt perchlorate produced, after isolation, rose-orange crystals $\text{Co}_2\text{L}_3(\text{ClO}_4)_4$, whose colour is typical of octahedrally coordinated cobalt. X-ray crystallography confirmed that the complex's structure is that shown in the colour-coded, space-filling model shown above, in which a (blue-coloured) cobalt ion can be seen peeking out between the ligands. As might be expected, strain in the ligand prevents the coordination of the cobalt ions from being perfectly octahedral. (Figures courtesy of Alan F. Williams.)

R.P.

Crowley and Baum, have been working on this problem using various approaches. One intriguing new idea is based on the abundance and isotope compositions of carbonates in ancient soils¹¹: higher soil carbonate abundances and more negative $\delta^{13}\text{C}$ values should correspond to higher atmospheric CO_2 . (Photosynthetic organisms produce isotopically light carbon; high ambient CO_2 partial pressures dilute this effect and produce isotopically heavier carbon.) A variation of this concept has recently been applied to the late Ordovician¹², the inferred atmospheric CO_2 level is about 16 times the present level, in remarkable agreement with Berner's model. Readers are left to judge for themselves whether this agreement is accidental.

Other than ice-core analyses, none of the proposed CO_2 palaeobarometers has gained widespread acceptance, possibly because all seem to involve large uncertainties. The experience with ice cores over the past decade, however, shows

that scientific disbelief will be suspended if a method produces self-consistent, reproducible results. One can only hope that one of the techniques being studied will eventually pass this test. If this happens, we will then be in a much better position to understand the intricacies of palaeoclimate. □

James F. Kasting is in the Department of Geosciences, The Pennsylvania State University, University Park, Pennsylvania 16802, USA.

1. Sagan, C. & Mullen, G. *Science* **177**, 52–56 (1972).
2. Owen, T., Cess, R. D. & Ramanathan, V. *Nature* **277**, 640–642 (1979).
3. Walker, J. C. G., Hays, P. B. & Kasting, J. F. *J. geophys. Res.* **86**, 9776–9782 (1981).
4. Crowley, T. J. & Baum, S. K. *J. geophys. Res.* **96**, 22597–22610 (1991).
5. Berner, R. A. *Science* **249**, 1382–1386 (1990).
6. Gaffin, S. *Am. J. Sci.* **287**, 596–611 (1987).
7. Volk, T. *Am. J. Sci.* **287**, 763–779 (1987).
8. Schwartzman, D. W. & Volk, T. *Nature* **340**, 457–460 (1989).
9. Kiehl, J. T. & Dickinson, R. E. *J. geophys. Res.* **92**, 2991–2998 (1987).
10. Crowley, T. J., Mengel, J. G. & Short, D. A. *Nature* **329**, 803–807 (1987).
11. Cerling, T. E. *Am. J. Sci.* **291**, 387–400 (1991).
12. Yapp, C. J. & Poeths, H. *Nature* **355**, 342–344 (1992).

Old bones match old stones

Bernard Wood

FROM their reappraisal of an 'old' fossil, together with a 'new' date for a little-known Kenyan fossil locality, A. Hill and colleagues (page 719 of this issue¹) claim that our own genus, *Homo*, may be 2.4 million years old. This is some half-a-million years older than the remains from Olduvai Gorge and Koobi Fora, which were previously regarded as the earliest fossil evidence for *Homo*.

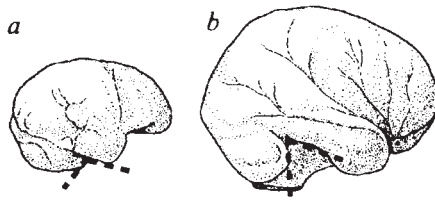
This is not the first proposal that the antiquity of *Homo* may stretch well beyond two million years (Myr). Two decades ago it was suggested that a cranium, KNM-ER 1470, from Koobi Fora, Kenya, together with several post-cranial remains of similar age, may be as old as 2.9 Myr². The fossil evidence has stood the test of time, for it is still either allocated to *H. habilis*, or is regarded as the type specimen of a new taxon, *H. rudolfensis*³; but the date has not. Early attempts to date the overlying tuff (volcanic ash layer) at Koobi Fora by conventional K-Ar isotope techniques suggested a tuff age of 2.6 Myr⁴, and subsequent ⁴⁰Ar/³⁹Ar total fusion analyses provided an age of 2.48 Myr. But in other hands, both K-Ar (ref. 5) and ⁴⁰Ar/³⁹Ar analysis of unaltered phenocrysts gave evidence of a younger date for the tuff of 1.88 Myr old; there is thus no unambiguous evidence for *Homo* from Koobi Fora older than 2 Myr.

Elsewhere the evidence for any greater antiquity for *Homo* is equivocal. The Shungura Formation to the west of the Omo river in southern Ethiopia preserves a fossil hominid record which extends back to around 3.0 Myr ago. The crown patterns of mandibular milk molars are quite distinctive in early hominids, and two such teeth from a mandible, Omo 222-2744, have convinced at least one experienced observer that this 2.1-Myr-old specimen should be allocated to *Homo*⁶. Twelve isolated teeth from Shungura Formation sediments, dated to 2.4 Myr old, have also been allocated to *Homo*, but the evidence is even less secure than for the milk teeth.

As regards southern Africa, claims for the presence of early *Homo* in the member 4 breccia of the cave site of Sterkfontein, which would indicate an age in excess of 2.5 Myr, rest on interpretations of the cranium, Sts 19. Some authors have stressed the resemblances of its petrous bone morphology to that of later *Homo*, but a subsequent study has concluded that it can be accommodated within the range of variation of *Australopithecus africanus*, the australopithecine that predominates in the member 4 hominid fossil sample.

The proposal of Hill *et al.*¹ that *Homo* may be at least as old as the evidence for stone artefacts is based on a reassessment of a hominid right temporal bone fragment, KNM-BC1, recovered from the Baringo region of Kenya in 1965. The fragment was found on the surface but matrix attached to it matched the lower calcareous grit lenses of the Upper Fish component of the Chemeron Beds, upon which the fossil was lying⁷.

Tobias, who in 1967 initially assessed the fragment⁷, considered the evidence it provided about the anatomy of the



The petrous crest is the ridge of bone dividing the cranial fossa that accommodates the temporal lobe of the brain from the cranial fossa that lodges the cerebellum. It is more open-angled in the apes and australopithecines and more acutely angled, and thus sharper, in *Homo*. *a*, *Australopithecus africanus*; *b*, *Homo sapiens*. (Reproduced from L. Aiello and C. Dean's *An Introduction To Human Evolutionary Anatomy*; Academic, London, 1990.)

mandibular fossa, the jaw joint, and the immediately surrounding regions. In 1967, there were relatively few East African hominids with which to compare the Chemeron fragment. Tobias acknowledged that the wide and relatively shallow mandibular fossa, the form of its boundaries and the nature of the mastoid process were australopithecine-like, but also conceded that a good proportion of the features were compatible with those seen in early *Homo*. The initial study confirmed its hominid status but did not assign it to *Australopithecus*. Others subsequently did so, however, partly, I suspect, because of the age of the Chemeron Beds, then estimated to be at least 4 Myr old.

The potential of the Baringo Basin as a source of evidence about Plio-Pleistocene hominid evolution remained unexploited until a joint initiative by Yale University and the National Museums of Kenya to establish the Baringo Palaeontological Research Project. It was soon realized that the Chemeron Formation not only includes outcrops dated to between 4.5 and 5.0 Myr ago, which have yielded the Tabarin mandible⁸, but much younger localities, little more than 2 Myr old, of which the Chemeron hominid site was one. New ⁴⁰Ar/³⁹Ar dates, from a tuff exposed

2.5 m below where the temporal fragment is believed to have come from, imply that the hominid dates to around 2.4 Myr ago, and the authors wisely decided to reassess the taxonomic affinities of the Chemeron hominid by exploiting the much richer comparative fossil sample assembled over the 25 years since its initial description. They assessed the distributions of ten character states and concluded that two, a sharp petrous crest (see figure) and the relatively medial position of the mandibular fossa, are shared between KNM-BC1 and *Homo* yet are not found in either *Australopithecus* or *Paranthropus*. It is on this basis that they allocate the Chemeron fragment to *Homo*. Their observations match previous suggestions that a relatively wide cranial base is found only in *Homo*⁹.

Fragmentary though it is, the Chemeron temporal preserves evidence that apparently discriminates between early *Homo* taxa (for review see ref. 10, to be published next week). The shallowness of the mandibular fossa virtually precludes its allocation to *Homo ergaster* and makes its allocation to *Homo habilis sensu stricto* unlikely. The Chemeron fragment probably represents the earliest evidence for *Homo rudolfensis*, and several subtle, but significant, features of the morphology of the cranial base suggest that by 2.4 Myr ago at least one hominid species was showing signs of the cranial remodelling that we associate with the genus *Homo*.

What, if any, are the phylogenetic implications of extending the temporal range of *Homo* back to at least 2.4 Myr ago? It does not prejudice the claims made by many that *Australopithecus afarensis* is the likely common ancestor of all later hominids. The primitiveness of *A. afarensis* has yet to be systematically and comprehensively tested, but Hill *et al.*'s analysis of a relatively narrow set of cranial characters suggests that it would make a plausible intermediate between hominids and the presumed common ancestor that hominids shared with the African apes.

The earlier date does coincide with what many concede is the earliest sound evidence for *Paranthropus boisei*, namely dental remains from the Omo Shungura Formation. The hominid fossil record between 2 and 3 Myr ago is too

- Hill, A., Ward, S., Deino, A., Curtis, G. & Drake, R. *Nature* **355**, 719-722 (1992).
- Leakey, R. E. F. *Nature* **242**, 447-450 (1973).
- Alexeev, V. P. *The Origin Of The Human Race* (Progress Publishers, Moscow, 1986).
- Fitch, F. J. & Miller, J. A. *Nature* **226**, 226-228 (1970).
- McDougall, I. *Nature* **294**, 120-124 (1981).
- Grine, F. *Cour. Forsch. Inst. Senckenberg* **69**, 69 (1984).
- Martyn, J. & Tobias, P. *Nature* **215**, 476-480 (1967).
- Hill, A. *Nature* **315**, 222-224 (1985).
- Dean, M. C. & Wood, B. A. *Am. J. phys. Anthropol.* **69**, 157-174 (1982).
- Wood, B. A. *Nature* (in the press).

sparse to read too much into this coincidence of 'first appearances', but 2.4 Myr ago is tantalizingly close to the 2.5-Myr-old date for the shift towards drier environments. The period between 2.5 and 2.3 Myr ago is marked by many extinctions and first appearances of species in the vertebrate fossil record; it is tempting to link these with the inferred changes in climate.

Palaeoanthropologists are beginning to develop explicit taxonomic and phylogenetic hypotheses about the early stages of the evolution of the genus *Homo*¹⁰. What is sorely needed, however, is more palaeontological evidence from the interval between 2 and 3 Myr

POPULATION GENETICS

Shooting the RAPDs

Phil Hedrick

MOLECULAR techniques have brought unimaginable power to analysis of the genetic relationships between organisms. But which technique is best for which question in, say, evolutionary or conservation studies, is a matter for debate, as was plain at a meeting last month*. Captive condors, confiscated salmon and cassava cultivars all came under discussion by a happily heterogeneous group of molecular and population biologists, and statisticians.

Some at the meeting preferred the traditional approach of allozymes, because the enzymes' functions and inheritance patterns are generally known and because the technique is cheap. Others favoured DNA fingerprinting based on VNTRs (variable numbers of tandem repeats) because all the data come out neatly arranged on one gel (for multi-locus probes) and there are large amounts of variability in nearly all organisms. But it was a new approach called random amplified polymorphic DNA (RAPDs, pronounced 'rapids') which created the biggest stir.

"If you like allozymes, you'll love RAPDs", announced R. Sederoff (Carolina State University), the chief cheerleader for RAPDs. Using the method, his laboratory constructed a genetic map of loblolly pine with 200 markers in 60 days, an order of magnitude faster than previous approaches. Improvements in the technique should cut this time to two weeks. The RAPD method uses a synthetic, ten-base-long chain to search for variations in DNA. Sederoff employs more than 700 different versions of the chains in his work, which requires only traces of DNA (as little as 1 nanogram). Between 40 and 80 per cent of these

ago. In the Omo region, Shungura and Koobi Fora span that period, but the predominance of riverine depositional environments in the former and the frustrating hiatus in the fossil record at the latter probably preclude their making a major contribution; West Turkana is a more likely source of evidence. Perhaps the Baringo Basin may yet provide the material that will allow researchers to test taxonomic hypotheses by matching them against hard fossil evidence. □

Bernard Wood is in the Hominid Palaeontology Research Group, Department of Human Anatomy and Cell Biology, University of Liverpool, Liverpool L69 3BX, UK.

sequences can give polymorphic markers, depending on the techniques used (J. Carlson, University of British Columbia). R. Vilgalys (Duke University) showed how RAPD markers could be used in oyster mushrooms to determine what is an individual, parentage and levels of gene flow.

RAPDs do have their drawbacks. For one thing, no one understands the exact basis of the variation the technique uncovers — it may be due to variation in sequences, or to insertions or deletions which cause loss of bands. This question may be resolved as soon as investigators at du Pont, represented at the meeting, finish sequencing DNA from some bands and come up with information that will allow them to probe and sequence DNA in and around the recognized ten-base sequence. Repeatability, expense and dominance are other potential snags. Two possible ways around the dominance problem (the absence of a band cannot be seen in heterozygotes) are to obtain many markers, thereby increasing statistical power, or to mark variants with codominant restriction sites. And to make the process cheaper, people could purify *Taq* polymerase, the expensive ingredient, in their own laboratories (information on how to do this is in the public domain). Those converted to the method claim that, with practice and care, repeatability is actually high.

Meanwhile, the applications of mol-

ecular genetic analysis are growing. DNA fingerprinting of the California condor has revealed important information on the relatedness among the remaining birds (O. A. Ryder, San Diego Zoo). Fifty condors are now in captivity, and two were released in southern California last month. Fingerprinting, combined with a multidimensional scaling analysis (C. Geyer and E. Thompson, University of Washington, Seattle) shows that the 52 condors divide into three distinct ancestral groups. These data could be crucial in formulating a successful reintroduction plan that adequately takes account of the existing condor gene pool.

Genetic analysis has also been used to try to sort out the sources of harvested salmon, but so far no technique seems to be the magic molecular bullet that can explicitly differentiate Canadian, US or Russian salmon (R. S. Waples, National Marine & Fisheries Service; P. Smouse, Rutgers University). (RAPD analysis is however planned.) Good allozyme data in North America have failed to distinguish between many stocks, while poor data from Russia do not allow a good genetic description of those fish. Analysis of the fish from an impounded



The California condor — two are now back in the wild.

Taiwanese ship with a \$14-million cargo of salmon suggests that 70 per cent came from an outgroup, maybe Siberia, but definitely not from Washington state as initially hoped by Washington researchers. Biology may in part explain the difficulty in finding good genetic markers. The separation of salmon into different stocks may have occurred only

*Molecular and Mathematical Approaches in Genealogical Analysis and Strain Identification Rutgers University, 10–11 January 1992.

recently, and leakage (imperfect homing) between stocks may make the perfect genetic marker rare or nonexistent.

Attempts to identify cultivars of apples, blackberries and cassava may meet with more success (B. A. Schaal, Washington University, St Louis). Separation of cultivars was indeed possible using multilocus DNA profiling in all these plants, except for some new sports from red-delicious apple cultivars. Of course, in this case the genetic difference may be only a gene (or a few genes) that influences the apple phenotype, and the time necessary to accumulate VNTR

differences may have been too short.

M. Lynch (University of Oregon) intends to extend the analytical approaches he developed for DNA fingerprinting to RAPDs, and the du Pont group are both automating the RAPD technology and developing new software to handle the large data sets that result. My guess is that answers to puzzles in population biology will now develop, well, RAPDly. □

Phil Hedrick is in the Department of Biology and Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, Pennsylvania 16802, USA.

X-RAY BINARIES

Optical signature of Cygnus X-3

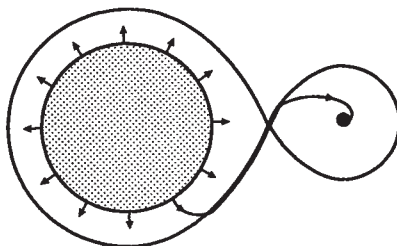
Peter S. Conti

ONE of the brightest X-ray sources in the sky, Cygnus X-3, has confounded explanation since it was discovered in 1966. The X-rays, produced as material is spewed from a neighbouring star onto the surface of a neutron star, have been accompanied by radio outbursts and by the most energetic γ -rays yet detected. Until now, it has been impossible to see the optical spectrum from the source, because of obscuration by dust along the line of sight. Hence the observation of near-infrared emission from the object, reported by van Kerkwijk *et al.* on page 703 of this issue¹, is of great interest. Furthermore, the authors find that the companion star is in the very last stages of evolution.

Cygnus X-3 is considered to be at a distance of 10 kiloparsecs, on the periphery of our Galaxy, given the strength of absorption of 21-cm radio waves along the line of sight. Its two components orbit one another with the extremely short period of 4.8 hours. And the γ -ray flux has a pulsation period of 12 ms, which has not, however, been confirmed at X-ray wavelengths². Galactic X-ray sources have their origin in the accretion of matter onto a neutron star, or a black hole, from an orbiting binary star; both high-mass and low-mass companion stars have been recognized.

Cygnus X-3 would normally be considered a low-mass binary because of its very short orbital period³. But van Kerkwijk *et al.*¹ find broad emission lines due to neutral and ionized helium in the object's 1- and 3- μ m spectra; hydrogen emission is conspicuously absent. Such spectral features are similar to those found in Wolf-Rayet stars, whose emission-line spectra are due to strong stellar winds, the nature of which is not yet fully understood. Wolf-Rayet stars are, for the most part, the luminous, highly evolved, helium-burning descendants of the most massive stars.

From its spectral similarity to nearby Wolf-Rayet stars, van Kerkwijk *et al.* suggest that Cyg X-3 is a high-mass X-ray binary with a Wolf-Rayet star of WN subtype (that is, having strong nitrogen emission). A well-known scheme⁴ predicts that a hot, massive O star in a binary X-ray system should end its life as a helium-burning Wolf-Rayet star, re-



The Roche lobe of a binary star defines a gravitational equipotential around the two components. At the midpoint, the gravitational pull of each is equal. Material reaching the lobe surface from one component will flow over the lobe and be accreted onto the surface of the other, causing bright X-ray emission in systems like Cyg X-3.

maining in orbit with its collapsed companion. The evolved system would persist as a strong X-ray source. Cyg X-3 may represent just this kind of binary. The source of accretion onto the neutron star would be the strong wind from the Wolf-Rayet component.

This interpretation is not without its difficulties, however. First, there is the short orbital period of the two partners, which is unprecedented for massive binary systems, of both O and Wolf-Rayet types. Such a short period could arise if the binary had undergone a period of extreme loss of mass and angular momentum, allowing the two partners to spiral towards one another. But more problematic is the comparative size of the Wolf-Rayet star and its Roche lobe (crudely, its gravitational sphere of influence; see figure) in such a tight binary.

van Kerkwijk *et al.*¹ point out that the Roche lobe of a putative 10-solar-mass WN star (for the high-mass case) in a 4.8-hour orbit with a 1.4-solar-mass neutron star would be 1.7 solar radii; they note that stellar-structure models of pure helium stars give radii of 0.9 solar radii, fitting comfortably inside the lobe (as shown in the figure). On the other hand, wind models of WN stars, which have been favourably compared with extensive observational data on nearby stars⁵, lead to 'core' radii (where the expansion velocity is negligible) that are three or more solar radii, larger than the Roche lobe. The wind of the WN star extends well outside this radius, so that the collapsed component of the binary would be completely buried in the wind of the other. Here, the star would be larger than its Roche lobe, a physically unrealistic situation. It might be that the velocity structure of Wolf-Rayet winds is not yet well modelled and Cyg X-3 might provide astronomers with a completely new observational constraint on the stellar wind of massive stars.

Could Cyg X-3 be a low-mass binary?

A mass of, say, 1 solar mass for the WN object would result in a much fainter and smaller object, comfortably fitting well inside its Roche lobe. Curiously, there are no models of low-mass WN stars as only two examples are known, both in binary systems. One of these is the nova-like variable, V Sagittae, which has an orbital period of 12.3 hours and comprises a 0.7-solar-mass WN star and a 2.8-solar-mass B-type subdwarf⁶. The other is AG Pegasus⁶, a 1-solar-mass WN star orbiting a 3-solar-mass M giant, with a period of 820 days⁷. Although neither of these systems is likely to evolve into anything resembling Cyg X-3, their optical spectra are well-nigh identical to those of high-mass WN stars.

Now that Cyg X-3 has been spectroscopically identified, explanations of the radio bursts and high energies generated in the system will need to be re-examined. Additional observations over the orbital cycle may be able to distinguish between the high- and low-mass interpretations for Cyg X-3. A flurry of additional papers can be expected now and the enigmatic nature of this close X-ray binary may be at hand. □

Peter S. Conti is at the Joint Institute for Laboratory Astrophysics, University of Colorado, Boulder, Colorado 80309, USA.

1. van Kerkwijk, M. H. *et al.* *Nature* **355**, 703–705 (1992).
2. Kitamoto, S., Mizobuchi, S., Yamashita, K. & Nakamura, H. *Astrophys. J.* **384**, 263 (1992).
3. Mitra, A. *Astrophys. J.* **384**, 255–262 (1992).
4. van den Heuvel, E. P. J. & de Loore, C. *Astr. Astrophys.* **25**, 387–395 (1973).
5. Schmutz, W., Hamann, W. R. & Wessolowski, U. *Astr. Astrophys.* **210**, 236–248 (1989).
6. Herbig, G. H., Preston, G. W., Smak, J. & Paczynski, B. *Astrophys. J.* **141**, 617–648 (1965).
7. Hutchings, J. B., Cowley, A. P. & Redman, R. O. *Astrophys. J.* **201**, 404–412 (1975).

PLATE TECTONICS

Geodesy tracks plate motion

Kenneth W. Hudnut

OUR current understanding of the large-scale movements of the Earth's great crustal plates is based mainly upon offsets of old geological features, on land as well as on the ocean floor. To what extent are models of relative plate motion, based on these aged offsets, applicable to today's plate motions? Discrepancies between the models and

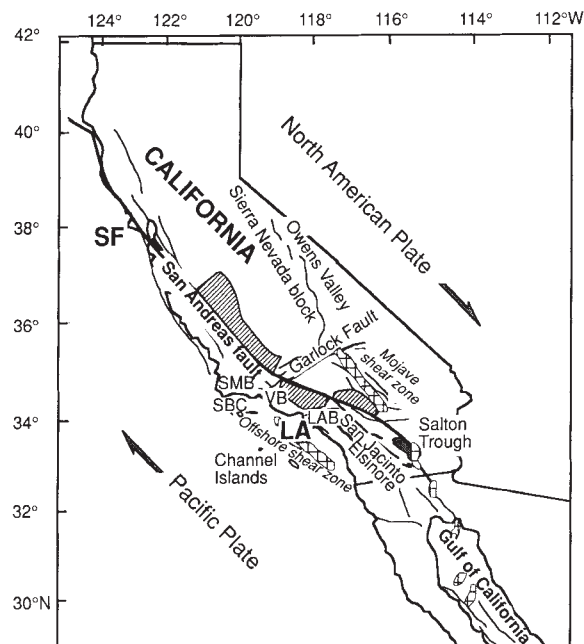
millimetres per year, allowing discernment of slip partitioning between the main and secondary faults and folds. Geodesists gathered recently at the meeting of the American Geophysical Union* in San Francisco to discuss the latest findings on plate motions in California as elucidated by GPS.

The challenge made to geodesists a few years ago was that unless they could pin down rates of less than 0.1 mm a year on individual active faults, they would have relatively little to contribute to research on active tectonics. Experience has since shown that whereas geological studies are particularly good for measuring slip rates on surficial strike-slip faults, geodetic studies can add much by resolving distributed motion across complex fault zones or areas of buried faults underlying folds at the surface.

The San Andreas discrepancy is the difference, previously thought to be as much as 20 mm a year, between the total lateral movement of the North American and Pacific plates past one another, and the amount of slippage seen on the San Andreas fault and its main subparallel fault systems. The discrepancy was reduced by the revision of plate motions performed by DeMets *et al.* two years ago¹, which

showed the lateral movement of the North American and Pacific plates to be only 47 mm a year. Geodetic and geological results now agree that the summed slip rate of the San Andreas and its main subparallel fault strands is 35 mm a year along most of the plate boundary². Another 8 mm a year of slip occurs in the Basin and Range province, and is probably localized almost entirely in the newly discovered Mojave shear zone that has been shown to connect the southern San Andreas fault to the Owens Valley³ (see figure).

The remaining 4 mm a year of boundary-parallel shear must occur in the borderland, and this apparently occurs in several zones along-strike of the plate boundary. Because the azimuth of motion in the Basin and Range is oblique to the plate boundary, the remaining displacement thought to be



The relative motion of the Pacific and North American plates, 47 mm yr⁻¹, is distributed over the San Andreas fault system (35 mm yr⁻¹), the Mojave shear zone (8 mm yr⁻¹), the offshore shear zone in the south (4 mm yr⁻¹) and in block rotation and convergence further north. Hatched areas, major thrust-sheet systems abutting the San Andreas fault. LAB, Los Angeles Basin; VB, Ventura Basin; SBC, Santa Barbara Channel; SMB, Santa Maria Basin; SF, San Francisco; LA, Los Angeles; SD, San Diego.

geological or geodetic data are of particular interest as these may reveal much about processes of crustal deformation. By studying details of the present motion between plates, we stand to learn a lot about limitations of plate tectonic theory, especially in continental regions where we know that the plates do not behave rigidly. Rather, they deform internally in various ways. For the boundary between the North American and Pacific plates in California, satellite measurements have helped unravel the complicated pattern of distributed plate motion across a zone several hundred kilometres wide. With the new Global Positioning System (GPS) satellites and receivers, it has been demonstrated that this evolving pattern of deformation can be studied with a resolution of several

*American Geophysical Union 1991 Fall Meeting, San Francisco, 9–13 December 1991.

RÉSUMÉ

Sleep on it

THE physiology and metabolism of a hibernating animal tick over slowly and benignly at body temperatures barely above freezing. What are the signals that cause the machinery to enter the idle mode? N. Kondo and J. Kondo (*J. biol. Chem.* **267**, 473–478; 1992) have discovered that when the chipmunk prepares for hibernation, a set of four plasma proteins disappear from the circulation, to reappear just before the creature awakes. One of them bears a homology with α_1 -antitrypsin; the other three are related only to each other and share a collagenous sequence at the N terminus. All four associate in a disulphide-linked complex. The same proteins were found in the blood of the ground squirrel, which also hibernates, but not in the tree squirrel or in rats, which stay awake. One cannot even speculate how these molecules — the inverse, one might say, of stress (or heatshock) proteins — might operate.

Silicon sieve

THE unexpected answer to those trying to devise stable, porous inorganic films may be silica (W.F. Maier, M. Weidorn & H. O. Schramm *Angew. Chem. Int. Edn Engl.* **30**, 1509–1510; 1991). The authors' search in fact started with an attempt to understand catalysis at platinum surfaces. They were surprised to find that the catalytic properties persisted even with a thin coat of silica evaporated onto the metal, despite an absence of defects or fractures. The explanation they proposed at the time, that the film was peppered with microscopic pores which transported reactant gases to and from the surface, is now confirmed by absorption-isotherm studies that reveal the films' capacity to soak up nitrogen. The pore diameter, it seems, varies by only an ångström or so either side of 6 Å, making the films as selective as commonly used zeolite molecular sieves.

Rabies in Peru

RARE documentation of an outbreak of human rabies, which resulted in the death of 29 people over a four-month period in 1990, appears in this week's *Lancet* (A. Lopez *et al.* **339**, 408–412; 1992). The place was two villages in the high Amazonian jungle of Peru and the transmitting agents were vampire bats (*Desmodus rotundus*), the main feral reservoir of rabies virus in South America. Lopez *et al.* can only conjecture why so large a proportion of the population of the villages (5 per cent) should have died in the outbreak. They propose that protection for isolated populations exposed to rabid vampires will come not from human vaccination, but from simpler measures such as the thorough cleaning of bat bites with soap and use of mosquito nets.

occurring in the borderland should have a substantial component of convergence perpendicular to the San Andreas fault. In the south, a shear zone passes between the coast and Channel Islands, whereas further north, localized convergence occurs across the Los Angeles basin, the Santa Barbara Channel and Ventura basins and the Santa Maria basin, as well as thrust-sheet systems that abut the San Andreas fault. Between these basins, it has been suggested that rotating crustal blocks accommodate the shear strain⁴.

Ambiguity

At the meeting, a series of talks showed that significant shear strain is observed between islands of the California borderland, and rapid shortening occurs across basins within the borderland. Ambiguity that caused controversy over plate boundary slip distribution during the early 1980s has now been almost completely resolved. Although there are important details to be worked out, the main quandaries that formerly surrounded the San Andreas discrepancy have been resolved, most of them by geodetic studies.

Challenges remain, of course. Searches with continuously recording GPS receivers for possible temporal variation in strain might explain, for example, observed fluctuations in the seismicity rate. And the details of deformation along the plate boundary with denser networks of GPS receivers could reveal transitions in strain as the plate-bounding fault system widens, narrows and bends between the Gulf of California and Cape Mendocino. Can the geological evolution of this plate boundary be understood well enough to give us a full picture of why the distributed plate boundary appears as it does today? Can we use the surficial data to constrain patterns of deformation in the mid- and lower crust, and clarify our vision of how the plate boundary looks at depth? How might this plate boundary become reorganized in the future?

It also remains to use southern California as a testing ground to address some general problems in continental tectonics and earthquake mechanics. Palaeomagnetic data suggest that block rotation played a large part in early structuring of the plate boundary, particularly in the Transverse Ranges and Mojave Desert. It will be interesting to see whether present-day rotations of rigid crustal blocks can be distinguished from simple shear across the zone. It has been proposed that convergence across the San Andreas fault is partitioned, occurring as slip on orthogonal detachment and thrust faults instead of oblique slip on the San Andreas. Geodesy might throw light on this fundamental process.

With the known geometry of secondary faults, it is clear that slip on these secondary structures must influence the San Andreas fault. Geodetic studies should help model such fault interactions and elucidate how earthquake recurrence is modulated on the San Andreas fault. How does an elastic response of the crust at the low strain rates observed geodetically enter into these observations, and are elastic models inadequate to explain the data?

GPS should also help reveal significant movement on deep structures, for instance on blind thrust faults in the Los Angeles region, where it is being used to identify which portions of these faults are accumulating strain that could be released in earthquakes. This work will lead to better estimates of earthquake hazard, and will help to quantify the mechanics of large thrust sheets abutting strike-slip faults, so improving our models of the earthquake cycle, and our attempts at long- and intermediate-term earthquake forecasting.

Speed

The quicker seismologists can learn about movement before, during and after earthquakes, and of aseismic slip events, the better. At present, it takes days to analyse GPS data. With more dense GPS networks, continuous monitoring and automated analysis, it should be possible to reduce the delay to hours. Geodesists could then match the rapid response of seismologists to earthquakes. It might also be possible to recognize precursory changes to warn of impending earthquakes. Measurements of coseismic movement could help disaster relief workers identify where they are needed most, for example in finding where tilting and disruption might affect water supplies.

Unfortunately, the conflicting military interest of the US Department of Defense, which runs the GPS satellites, could deprive earth scientists of the system's greatest advantages. Planned restriction of real-time access to part of the GPS signals could hinder moves towards automated data processing. How this political dispute between civilian and military interests is resolved will be of great interest to those living in earthquake zones. □

Kenneth W. Hudnut is in the Seismological Laboratory, Division of Geological and Planetary Science, California Institute of Technology, Pasadena, California 91125, USA.

1. DeMets, C., Gordon, R. G., Argus, D. F. & Stein, S. *Geophys. J. Int.* **101**, 425–478 (1990).
2. Lisowski, M., Savage, J. C. & Prescott, W. H. *J. geophys. Res.* **96**, 8369–8389 (1991).
3. Savage, J. C., Lisowski, M. & Prescott, W. H. *Geophys. Res. Lett.* **17**, 2113–2116 (1990).
4. Jackson, J. & Molnar, P. *J. geophys. Res.* **95**, 22073–22087 (1990).

Born again babies

ALONE among mammals, human beings are born with great difficulty and danger. It's all the fault of the baby's big head. In evolutionary terms, a big-headed, big-brained baby is so valuable that it is well worth facing certain pain and possible death to bear it. Even so, Daedalus plans to reduce the risks.

He recalls the technique of giving birth to a baby under water. This is said to be more relaxing for the mother. The baby doesn't drown because it is still being nourished from the placenta and need not begin breathing until the umbilical cord is cut. It may not even realize that it has left the womb. Daedalus is now taking this idea to extremes. He is designing a sort of external auxiliary womb, a water-filled tank that can be attached to the mother-to-be by a short flexible pipe. At a stage when the baby's head is still small enough not to pose serious risks, birth can be induced by the usual drugs. The baby emerges into the tank, still attached to the mother, and can stay and grow there until it is ready to be 'born again' into the atmosphere.

An external womb will obviously be massively inconvenient. Even if supported on wheels it will severely limit the mother's mobility. The problems of maintaining sterility and hermetic sealing will complicate matters further. It may be best to design the whole thing as a sort of high-tech child bed in which the mother can lie and be looked after, for however many weeks it takes for the child to develop fully.

The benefits, however, will also be great. Slim-hipped mothers may well prefer a month or two in bed to the threat of caesarian section, while those at risk of premature birth could turn events to their own advantage. The baby could be inspected through a transparent panel, and minor damage from the birth process could be corrected via a glove insert. The mother could also touch and caress the baby; if the two could also look at each other they might become emotionally bonded even before final delivery.

Daedalus also comments that the normal nine-month pregnancy must be an evolutionary compromise between the interests of the mother and those of the baby. His artificial womb could relax this compromise, and allow a pregnancy of ten months or even longer: as long, in fact, as the placenta could deliver adequate nourishment. Finally, a far bigger and better-prepared child would emerge into the world. Unlike the rest of us, it should be free of the 'birth trauma' which, according to some psychologists, lurks worryingly in the normal human subconscious. Or it could have two of them.

DAVID JONES

Degassing Lake Nyos

SIR — Lakes in volcanic regions can absorb high concentrations of magmatic CO₂, and spontaneous degassing can occur if the pressure is reduced. On 21 August 1986, Lake Nyos in Cameroon released an estimated 0.12 km³ (240,000 tonnes) of CO₂ (ref. 1) which swept downhill and killed more than 1,700 people. Many specialists believe that the surface water, cooled by rain, became denser than a CO₂ rich layer and convective overturn occurred^{1,2}. Whatever started the movement, the rising water became supersaturated at reduced pressure and degassed. The gas/water mixture was very buoyant, and gas-driven circulation started, entraining deeper CO₂-rich layers, which also degassed, frothing over a 75-m-high promontory³. Input of CO₂ is estimated to be at least 10 m³ min⁻¹ (ref. 4) so the estimated loss in 1986 will be replaced in, say, 20 years. A method of controlled degassing of Lake Nyos and similar lakes is needed for the safety of the local population⁵. The remote nature of the lakes requires that the method is simple, cheap and maintenance-free.

The latent energy of dissolved gas could be used to power a gas-lift in a tube extending from near the lake bed to the surface. When upwards flow is started by a pump, effervescence within the tube would reduce the density of the fluid, and flow would continue passively, an elegantly simple solution. Some sediment in the water would nucleate bubbles and increase the efficiency of degassing. A 9-mm tube has been used analytically in this way⁶ and large-scale experiments are also being undertaken. If flow stops for lack of CO₂, the lake is safe. Flow can be controlled by tube diameter or a valve at the top. Semi-rigid polythene tubing of 10–20-cm diameter would be convenient for degassing the lake. The top will need to be firmly anchored to resist the jet propulsion effect.

The deep water contains about 6,000 mg kg⁻¹ CO₂, or about 3 l CO₂ at atmospheric pressure at 20 °C, and would become supersaturated at depths of less than 27 m. Bottom water may contain two or three times as much CO₂ (ref. 7). An input of 20 kg CO₂ min⁻¹ would be equalled by degassing or removing about 3,500 l water per min.

Given the volume increase to 14,000 l gas/water mixture at the surface, many tubes will be needed. Larger volumes with less CO₂ content must be degassed to reduce CO₂ to safer levels. Selective removal of the most buoyant upper part of the CO₂-rich water would also help to stabilize the lake.

The degassed water must be disposed of carefully to maintain the stability of the lake and avoid triggering another uncontrolled degassing. Most lakes have cold dense deep water and a buoyant warm surface layer. Lake Nyos, however, is heated geothermally, with warm, possibly CO₂-saturated water percolating in through the bottom⁷. The coldest water layer (about 22.5 °C) is at about 25 m, and temperature increases to around 24 °C at the lake bed (200 m). Surface temperature varies seasonally around 24 to 26 °C. The stability is almost entirely due to the mass of CO₂ dissolved in the bottom water increasing its density. A decrease in temperature of the CO₂-poor circulating surface layer by only 3.5 °C would cause it to sink below water containing 3,000 mg kg⁻¹ CO₂ at a depth of 100 m (ref. 1) and start massive degassing. Degassed water can only be added to the surface if it increases buoyancy there. This depends on extent of degassing, and reduction in temperature by exsolution and expansion of gas, and will only be determined experimentally on site. The gas-lift can raise water until the volume of water in the column above the lake surface almost equals the volume of gas below the surface, practically perhaps 10 m. On sunny days in the dry season, degassing water could be lifted onto the land and warmed by solar heating at it trickled back. The heat input to the surface would increase stability.

This stability problem can be overcome by discharging the water outside the lake, which also avoids problems of incomplete degassing. More than 7,500 tonnes of water per day need to be degassed, and normal outflow is 50,000 tonnes per day in August/September², the rainiest period, so lake levels can be maintained. Also, the less warm surface layer would be decanted from the spillway, so lake stability would increase further. At 50,000 tonnes a day it would take 10 years to flush the 175 million tonnes of water in the lake. Reducing the level of the lake will reduce its capacity for CO₂. Removing water from the bottom will not reduce pressure within the water column, but there are likely to be greater quantities of CO₂ in the sediments and rock beneath. This should degas readily because of the nucleating action of solids. It would be

advantageous in the long term to lower the lake level, but only at a time and rate such that degassing sediment will not upset the waters above.

CO₂ concentration in Lake Nyos has been increasing since the last limnic eruption, and must be released eventually. As soon as any layer becomes saturated, any disturbance will cause a degassing disaster. While the risks and effect of degassing must be monitored carefully, the risks of delay are clear.

HUGH FLETCHER

*School of Biology and Biochemistry,
The Queen's University of Belfast,
Belfast BT9 7BL, UK*

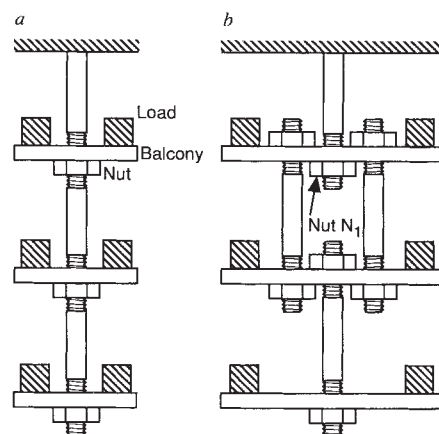
1. Gignenbach, W. F. J. *Volcanol. Geotherm. Res.* **42**, 337 (1990).
2. Kling, G. W. *Science* **137**, 1022 (1987).
3. Freeth, S. J. & Kay, R. L. F. *Nature* **325**, 104 (1987).
4. Freeth, S. J. *et al. Nature* **348** (1990).
5. Freeth, S. J. *J. Afr. earth Sci.* **13**, 411 (1991).
6. Halbwachs, M. *et al. Int. Wk Grp Crater lakes Newslett.* **3**, 8 (1990).
7. Nojiri, Y. *et al. Nature* **346**, 323 (1990).

Real-life failure

SIR — The paradoxical behaviour of a network of springs and strings reported by J. E. Cohen and P. Horowitz (*Nature* **352**, 699–701; 1991) recalls an analogous situation that resulted in considerable loss of life through the structural failure of a Kansas city hotel.

The true world is even more confusing. The diagram illustrates a paradoxical situation which resulted in considerable loss of life in the Kansas City hotel disaster.

A 'routine' engineering change during construction from *a* to *b* tripled the shear



load on nut N₁, causing three hotel balconies to fall when N₁ split. Note that *b* distributes the load over seven nuts instead of the three nuts of *a*. This structural failure, as in the case of the springs and electrical circuits described by Cohen and Horowitz, falls into the category of series-to-parallel transformations.

HOWARD I. PODELL

*28 Beachfront Lane,
New Rochelle,
New York 10805, USA*

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references. □

AIDS vaccine developments

SIR — In assessing vaccine strategies for preventing AIDS, Stott *et al.* reported¹ that macaques immunized with uninfected C8166 cells² (a T-cell line of human origin) were protected against subsequent intravenous infection by simian immunodeficiency virus (SIV), which induces an AIDS-like disease in macaques. Further investigation of the humoral response of vaccinated monkeys with SIV-infected and inactivated cells or with purified inactivated virus, revealed that all animals have antibodies directed against cellular components.

the strongest humoral immune response to both SIV and C8166 antigens. We also detected high ELISA titres against fresh human PBLs, whereas these animals raised weak antibody responses against rhesus macaque PBLs.

Eighteen weeks after the first challenge experiments, the protected animals of the high-dose vaccinated group were reboosted with the same vaccine preparation and challenged intravenously 2 weeks later with 10 AID₅₀ of an homologous SIV_{MAC}251 strain, produced by co-cultivating spleen cells of an infected

antibody crossreactivity between cellular and lentivirus components, which Maddox suggests³ could contribute to the understanding of AIDS pathogenesis. Nevertheless, our results show that protection of macaques cannot be related to an immune crossreactivity between cellular components and SIV antigens. Rather, we suggest that the strong anti-human cell immunity developed by our vaccinated macaques efficiently participates in protection (as previously suggested⁶) through cellular antigens on the envelope of the human PBL-grown viruses used for the first challenge. The results we report here demonstrate the absolute necessity of designing vaccine experiments where viruses used for challenge infection are grown in PBL compatible with the host species. In addition, an alternative strategy to the investigation of the anti-lentiviral specific immune responses can be provided by highly purified antigens whatever their origins: inactivated viruses, recombinant or synthetic subunits.

ROGER LE GRAND
BRUNO VASLIN
GUILLAUME VOGT
PIERRE ROQUES
MARC HUMBERT

DOMINIQUE DORMONT*

*Laboratoire de Neuropathologie
expérimentale et Neurovirologie,
Centre de Recherches du Service
de Santé des Armées,
Commissariat à l'Energie Atomique
BP6, 92265 Fontenay aux Roses Cedex,
France*

ANNE MARIE AUBERTIN

*Laboratoire de Virologie,
Université Louis Pasteur,
67000 Strasbourg, France*

* Also at Institut Pasteur, 75015 Paris, France.

ELISA TITRES AGAINST SIV AND CELLULAR ANTIGENS AT THE DAY OF CHALLENGE

Antigens detected	2 mg vaccinated				0.4 mg vaccinated			Control	
	8,715	8,738	8,744	8,770	8,758	8,762	8,778	8,771	8,783
SIV	5	5	5	5	5	4	4	Neg	Neg
C8166 cells	4	4	5	4	4	3	3	Neg	Neg
Human PBL	3	3	4	3	ND	ND	ND	Neg	Neg
Rhesus PBL	1	2	2	1	1	1	1	Neg	1

Neg, negative result. Regarding the high sequence homology of HIV-2 and SIV, the specific anti-SIV serum reactivity was determined⁴ using an HIV-2 antigen detection assay. We used the ELAVIA II kit (Diagnostics Pasteur) using as second antibody a peroxidase-labelled anti-monkey immunoglobulin G. Titres were determined as the last serum dilution, giving a significant positive optical density (OD) at 490 nm when compared to measures of 10 preimmune sera at the same serum dilution (cut off = OD average of preimmune sera + 2.26 s.d.). ND, not determined. ELISA titres expressed as log₁₀.

Moreover, protection of these macaques against an homologous SIV challenge correlates with the ELISA titres against C8166 cells. Both vaccine preparations and the challenge virus stock were obtained from the same culture system: SIV-infected C8166 human cells.

As part of the European Communities' AIDS Concerted Action programme, we immunized rhesus macaques with the same column-purified vaccine preparation (formalin-inactivated SIV_{MAC} strain 251, 32H isolate, grown on human C8166 cells, provided by M. Cranage and P. Greenaway) in alum adjuvant. Two groups of animals received either 2 or 0.4 mg immunogen (days 0, 30, 60 and 135) and were challenged, 2 weeks after the final boost, with 10 AID₅₀ (animal infectious dose 50%) of an heterologous SIV_{SM} virus stock (provided by P. Putkonen) produced on fresh human peripheral blood lymphocytes (PBL)^{3,4}. All four macaques in the high-dose immunized group, and two of three animals in the low-dose group, were protected more than 4 months after intravenous challenge.

We analysed the antibody responses of these monkeys. All animals exhibited high ELISA titres to uninfected C8166 cells as well as lentiviral-specific antigens. The high-dose vaccinated animals developed the highest humoral response to lentiviral and cellular antigens (see table). Surprisingly, ELISA titres were 10 times lower in the low-dose vaccinated group except for the unprotected monkey (No. 8,758), which developed

monkey with fresh macaque PBLs: this virus has therefore never been exposed to a human cell. All monkeys developed clinical and biological evidence of infection (virus isolation by coculture, PCR, anamnestic response) 15 days after the second challenge experiment.

Controversial results from vaccine experiments highlight the difficulty in interpreting such experiments: cautious conclusions should include detailed discussions about controls. Our experiments do not exclude the possibility of

SIR — Several groups have independently demonstrated that macaques may be protected from infection with simian immunodeficiency virus (SIV) by immunization with inactivated vaccines, based on either whole inactivated SIV or fixed SIV-infected cells⁷⁻⁹, raising expectations that the development of a vaccine for the protection from human immunodeficiency virus (HIV) infection in man may be successful. But it has become apparent that non-virus-specific cellular antigens present in SIV vaccine preparations used in those experiments may have played a role in the protection observed. The SIV virus stock used for challenge of vaccinated monkeys had been prepared from SIV-infected human T-cell lines, the same or similar to those used for production of the vaccine preparation itself¹.

We have carried out a vaccination challenge experiment in the same SIV-macaque model as part of the European Communities' Concerted Action pro-

gramme. Our results indicate that at least part of the protection induced by inactivated SIV preparations is not due to immunization with non-virus-specific human T-cell antigens. To this end we have compared the efficacy of two SIV whole virus vaccine preparations, administered to two groups of seven and eight rhesus monkeys (*Macaca mulatta*), respectively. The first vaccine was an inactivated whole SIV preparation adjuvanted with muramyl dipeptide (MDP) previously demonstrated to elicit protection; the second was an SIV-iscom preparation containing both the Gag and Env proteins of SIV. These vaccines had been prepared from the SIV_{MAC} strain 251 (32H), propagated on the human T-cell line C8166 as described by Stott *et al.*¹. Inactivated measles virus adjuvanted with MDP (MV-MDP) and an MV-iscom preparation¹⁰ served as controls, each of which were inoculated into two separate groups of four rhesus monkeys. Two weeks after the fourth intramuscu-

PROTECTION INDUCED BY SIV VACCINES IN RHESUS MACAQUES

Group	Type of vaccine	Type of i.v. SIV challenge (10 MID ₅₀)	Monkeys protected from viraemia*/monkeys per group	Serum antibody titres measured in ELISA (log ₁₀ ± s.d.) at day of challenge†		
				Antigen		
A	SIV-MDP	cell-free	4/4	SIV-env	C8166	RhPBMC
	SIV-iscom‡	cell-free	3/3	2.4±0.4	2.0±0.2	1.9±0.1
	MV-MDP	cell-free	0/2	3.4±0.2	2.7±0.2	2.7±0.2
	MV-iscom	cell-free	0/2	≤0.8	2.1±0.1	1.8±0.2
B	SIV-MDP	infected PBMC	2/4	≤0.8	1.9±0.1	1.9±0.3
	SIV-iscom	infected PBMC	2/4	3.3±0.4	2.1±0.1	1.9±0.1
	MV-MDP	infected PBMC	0/2	3.6±0.3	2.8±0.2	2.7±0.2
	MV-iscom	infected PBMC	0/2	≤0.8	≤1.5	1.9±0.3
				≤0.8	≤1.5	2.1±0.1

* Viraemia was demonstrated by co-cultivation of PBMC with C8166 cells and subsequent demonstration of SIV antigen in the medium by P27-antigen capture assay. Results were confirmed by showing serum antibody induction (MV-vaccinated monkeys) or booster reaction, in an SIV env-specific ELISA and Western blot assay.

† SIV-Env: inhibition of reactivity of labelled SIV-neutralizing mouse monoclonal antibody (KK5) by serial dilutions of monkey sera. KK5 kindly provided by Dr K. Kent (through MRC-ADP).

‡ One animal not included (died during recovery at day of challenge).

C8166/RhPBMC, reactivity of serial dilutions of monkey sera with solubilized membrane protein fraction (about 10 µg ml⁻¹) of respective cells attached to ConA-coated wells.

lar vaccination all the monkeys were challenged intravenously with either 10 monkey infectious doses (MID₅₀) SIV_{MAC}251 (32H) propagated in C8166 cells (group A), or with SIV-infected peripheral blood mononuclear cells (PBMC) obtained from an SIV-infected rhesus macaque (1×C) (group B). These PBMC had been prepared from heparinized blood of the monkey, 11 months after experimental infection with the same cell-free challenge stock of SIV_{MAC}251 (32H), frozen in 5×10⁵ cell aliquots and subsequently titrated *in vivo* in rhesus macaques. The equivalent of 10 MID₅₀ of these cells was used as challenge dose for group B.

The results are shown in the table. All the monkeys of group A vaccinated with SIV-MDP or SIV-iscom were protected from developing SIV viraemia for the 12-week observation period after intravenous cell-free SIV_{MAC}251 (32H) challenge, whereas all the MV-MDP and MV-iscom vaccinated monkeys developed SIV viraemia within 4 weeks after receiving the same cell-free challenge. Also all the MV-MDP and MV-iscom vaccinated monkeys of group B developed SIV viraemia within 2 weeks after intravenous challenge with SIV-infected PBMC. Two out of four SIV-MDP vaccinated monkeys and two out of four SIV-iscom vaccinated monkeys of group B were protected from SIV viraemia for the 9-week observation period after intravenous challenge with the SIV-infected PBMC.

These data were confirmed by demonstrating that all the monkeys that became SIV-viraemic also showed booster responses in their serum antibody titres measured in an SIV env-specific enzyme-

linked immunosorbent assay (data not shown). At the day of challenge, all SIV-vaccinated animals had developed serum antibody titres in this assay. In addition, antibodies directed against C8166 cells and rhesus PBMC present at that day were measured (see table). The protection found in group B among SIV-vaccinated monkeys did not correlate with the levels of serum antibody titres to C8166 cells or PBMC in these animals.

This is the first demonstration in the SIV-macaque model that vaccination can protect against challenge with cell-associated SIV, and is the first report that vaccinated, previously unchallenged non-human primates can be protected from infected PBMC from another infected animal. As the SIV challenge material used in group B was directly prepared from infected PBMC of the homologous species, the partial protection observed against this severe intravenous challenge should be attributed to immunization with SIV-specific antigens. Consequently the suggestion by Stott *et al.*¹, that virus- as well as cell-specific components can still be involved in protection after vaccination with inactivated SIV preparations, is strongly supported by our observations.

ALBERT OSTERHAUS

PETRA DE VRIES

National Institute of Public Health and
Environmental Protection,
PO Box 1, 3720 BA Bilthoven,
The Netherlands

JONATHAN HEENEY

ITRI-TNO,
PO Box 5815,
2280 HV Rijswijk,
The Netherlands

SIR — Stott *et al.*¹ have presented evidence that the protection elicited with an SIV vaccine⁹ containing fixed SIV-infected cells of human origin (C8166) may, at least in part, be mediated by anti-cell responses resulting from xenotransplantation. This surprising observation raises important questions about the nature of the protective responses elicited by other SIV vaccines.

As part of an AIDS vaccine development project, within the MRC AIDS Directed Programme and the European Communities' Concerted Action programme, we have immunized 10 rhesus macaques with a partially purified, inactivated virus vaccine based on the 32H isolate of SIV_{MAC} strain 251 grown in C8166 cells (manuscript in preparation). Four animals (group A) each received four 500-µg doses of vaccine given in the Syntex adjuvant formulation 1 (saf-1). Another two groups of three animals each received the same vaccine given in alum, either as four 500-µg doses (group B) or as four 100-µg doses (group C). All animals were challenged intravenously 2 weeks after the final vaccine boost with 10 median monkey infectious doses (MID₅₀) of cell-free virus. Group A were challenged with SIV_{SM}B670, a related but antigenically distinct strain of SIV, grown on human peripheral blood mononuclear cells (PBMC)⁸. Animals in groups B and C were challenged with homologous virus, that is the 32H isolate of SIV_{MAC}251 grown on C8166 cells.

Although unvaccinated control animals all became infected, all of the vaccinated animals were protected from infection as determined by the inability to recover virus from PBMC, inability to detect proviral DNA in PBMC and lack of an anamnestic antibody response. Animals were subsequently boosted further with vaccine, formulated as previously, at 16 months (group A) and 6.5 months (groups B and C) after initial challenge. At 2 weeks after the additional boost, all 10 animals plus 4 unvaccinated controls were challenged with 10 MID₅₀ of cell-free SIV_{MAC}251 grown in monkey PBMC. Virus was recovered from all of these animals as early as 2 weeks after challenge. Thus no protection from challenge was elicited despite the fact that the additional vaccination boosted the SIV-specific antibody response. In this experiment we have not

1. Stott, E. J. *et al. Nature* **353**, 393 (1991).
2. Salahuddin, S. Z. *et al. Virology* **129**, 51–64 (1983).
3. Fultz, P. V. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 5286–5290 (1986).
4. Putkonen, P. *et al. Nature* **352**, 436–438 (1991).
5. Maddox, J. *Nature* **353**, 297 (1991).
6. Schwartz, D. H. *Nature* **354**, 439 (1991).
7. Desrosiers, R. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 6353–6357 (1989).
8. Murphy-Corb, M. *et al. Science* **246**, 1293 (1989).
9. Stott, E. J. *et al. Lancet* **336**, 1538–1541 (1990).
10. Morein, B. *et al. Nature* **308**, 457–460 (1984).

formally established if the animals would have been protected from rechallenge with SIV grown in human cells, but in previous experiments we have obtained protection following reboosting.

At first sight our results suggest that the challenge viruses used in these experiments may have acquired properties, either by genetic selection or, perhaps more likely, epigenetically, from the cells in which they were grown. Thus anti-cellular responses elicited by immunization may have been responsible for the protection observed when both vaccine and challenge virus were derived from the xenogeneic human cells. But caution is necessary in interpreting these results: because the virus used for rechallenge was produced in rhesus monkey PBMC from the original SIV_{MAC251} stock, it could have contained variants that were critically divergent from the cognate virus used to prepare the vaccine.

It is indeed remarkable that crude inactivated SIV vaccines provide complete protection against challenge with various antigenically distinct viruses which, in unvaccinated animals, invariably induce AIDS. However, this failure to protect against virus grown in monkey PBMC and the significance of anti-cellular responses must be understood so that the protecting mechanisms can be defined and applied to AIDS vaccine development.

M. P. CRANAGE
L. A. E. ASHWORTH
P. J. GREENAWAY

PHLS Centre for Applied Microbiology
and Research,
Porton Down,
Salisbury SP4 0JG, UK

M. MURPHEY-CORB
Delta Regional Primate Research
Center,
Tulane University,
Covington,
Louisiana 70434, USA

R. C. DESROSIERS
New England Regional Primate
Research Center,
One Pine Drive,
Southborough,
Massachusetts 01772, USA

Temperature time-series?

SIR — Ghil and Vautard¹ applied the method of singular-spectrum analysis to a 135-year time-series of global annual-average near-surface temperatures, and found a pair of high-variance empirical orthogonal functions (EOFs) characterizing a temperature oscillation with a period of slightly more than 20 years. Elsner and Tsonis² repeated the analysis considering the most recent 130-, 110-

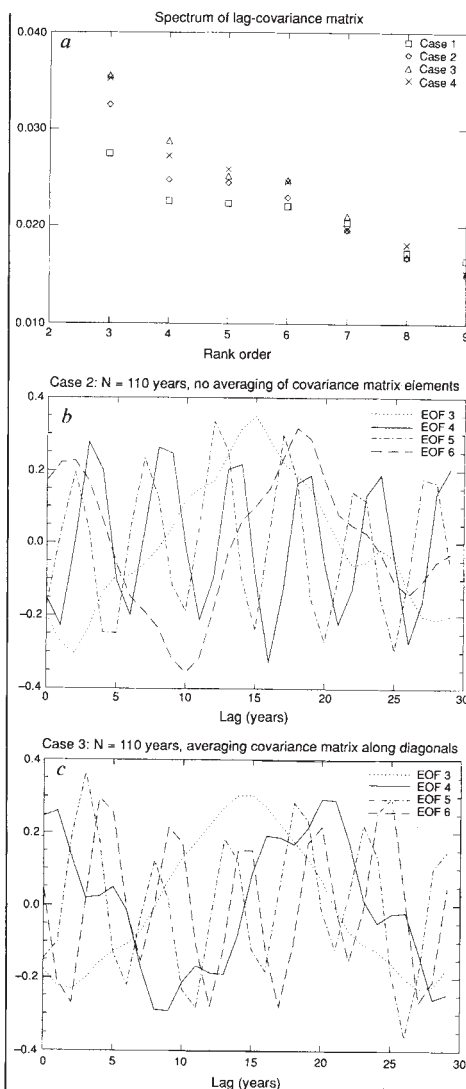
a, Portion of spectrum of covariance matrix constructed^{1,2} from $M = 30$ lagged copies of the IPCC consensus time-series of global temperatures³. Note the linear scale, to emphasize the similarity of eigenvalues 4, 5 and 6 in all cases considered. Ordinates: a, eigenvalue; b, c, eigenvector. b, EOFs 3–6 for case 2: $N = 110$ years, no averaging of covariance matrix elements. Note EOFs 3 (dotted line) and 6 (dashed line) are in quadrature, indicating a 20–30-year oscillation, while EOFs 4 (solid line) and 5 (dash-dot line) indicate a 5–6 year oscillation. c, EOFs 3–6 for case 2: $N = 110$ years, averaging covariance matrix elements along diagonals. Note EOFs 3 and 4 are now in quadrature.

and 90-year segments of the time-series. They report that the EOF 'signature' of the 20-year oscillation appears only if all 130 years are included, suggesting it may be an artefact of the (unreliable) early part of the time-series.

We found this pair of results intriguing, because Ghil and Vautard also show a reconstruction of the original data by projection onto a number of subsets of their EOF eigenbasis, including one in which only the lowest 4 EOFs (which include the 20-year oscillatory pair) are retained (their Fig. 3a). The amplitude of the 20-year oscillation does not appear to diminish in the later part of the time-series, as would be expected if the conclusion of ref. 2 (that all the power in this pair of EOFs is confined to the earlier part) were correct. Using the IPCC consensus time-series³ (data kindly provided by D. E. Parker), we reproduced both sets of results. Are they a paradox?

The answer is no. If we consider the relevant portion of the spectrum of the lag-covariance matrix, including all 130 years (a in the figure, case 1), we notice that the eigenvalues of EOFs 4, 5 and 6 are remarkably similar. A small change in analysis technique or time-series length might therefore change the order in which these EOFs appear. Sure enough, if we consider only the last 110 years (a in the figure, case 2), and inspect EOFs 3 to 6 (b), we notice that EOF 6 (dashed line) is in quadrature with EOF 3 (dotted line), forming a 20–30-year oscillatory pair, whereas with the full 130-year time-series, it is EOFs 3 and 4 which are in quadrature (Fig. 1b of ref. 1). Thus, removing the first 20 years of the time-series causes EOFs 4 and 6 to exchange positions in the eigenvalue rank-ordering, through a very slight change in their respective eigenvalues.

Ghil and Vautard impose a Toeplitz structure on the lag-covariance matrix. We are unsure if this is advisable, as it imposes symmetries on the problem which may not be present in a short data series. But if we do the same (a, case 3), we can make EOF 4 and EOF 6 swap



back again (c). If we consider only the last 90 years of the series, they re-exchange. If the rank order of these EOFs is so sensitive to details of the analysis, the order itself cannot have any physical significance. In dealing with such a short time-series, the possibility that results may be a consequence of a few anomalous decades, or an artefact of time-series and window lengths, must always be considered carefully. But Elsner and Tsonis have definitely failed to prove their case against the 20-year oscillation.

MYLES R. ALLEN
PETER L. READ

Department of Physics,
University of Oxford,
Parks Rd,
Oxford OX1 3PU, UK

LEONARD A. SMITH
Department of Engineering,
University of Warwick,
Coventry CV4 7AL, UK

- Ghil, M. & Vautard, R. *Nature* **350**, 324–327 (1991).
- Elsner, J. B. & Tsonis, A. A. *Nature* **353**, 551–553 (1991).
- Houghton, et al. *Climate Change: the IPCC Scientific Assessment* (Cambridge University Press, 1990).

Systematic progress

Niles Eldredge

Reading the Shape of Nature: Comparative Zoology at the Agassiz Museum. By Mary P. Winsor. University of Chicago Press: 1991. Pp. 324. \$21.95, £17.50 (pbk); \$49.95, £39.95 (hbk).

THE health of natural history museums reflects to a great degree the esteem felt for systematic biology at any given moment. Despite dismal news of fiscal travail in many of these institutions, hope continues to spring afresh for them: in a world just awakening to the ghastly rate of shrinkage of organic diversity, it is rapidly becoming appreciated that museums are libraries of diversity, housing not only the specimens but the only people we can turn to when we ask, "what is out there?", "how fast are species disappearing?" and perhaps even, "what can we do to save them?". Systematics is gaining renewed importance in both scientific and world affairs, which bodes well for natural history museums.

Thus a study of the founding of one particular natural history museum and the early history of systematics research there has welcome relevance to the modern world. *Reading the Shape of Nature* is, by Winsor's own reckoning, neither a history of systematics nor an overly detailed account of the origin and early development (to the 1930s) of Harvard University's Museum of Comparative Zoology. Rather, in portraying the pursuit of systematics at that institution, Winsor intends to provide a microcosm of the progress of the science generally.

Winsor is at her best in describing the infighting at the museum between its founder, Louis Agassiz, staunch opponent of evolution, and his disciples and younger colleagues. Agassiz was a man of considerable ability, opinion and reputation. While he is on the scene (he died in 1873), Winsor fulfils her ambitious goal of melding social history within the confines of a single institution with the broader sweep of intellectual history. We see, in perhaps the book's most arresting section, a microcosm even within the museum's microcosm, as suc-

cessive members of staff grapple with crayfish systematics in the years 1867–85. All the while, evolution was becoming increasingly respectable, and, mirroring work in the world at large, an awareness of it became increasingly present in the published work of the museum's systematists, as Winsor nicely points out.

But then Agassiz dies. Halfway through the book he is replaced by his son Alexander, a man who accepts evolution but has little real intellectual

done' as are the major ones, even when the latter are wrong (as Agassiz père was in opposing Darwin). Glimpses of issues become, in the jargon of modern systematics, 'paraphyletic': we no longer get to see the main action, but rather see only disconnected bits and pieces, without the context of overall extramural opinion.

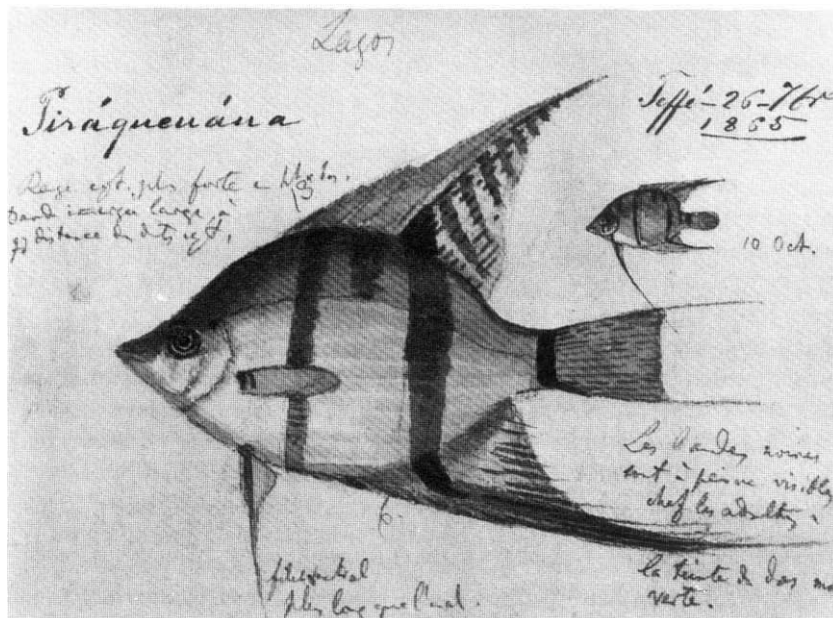
Scientific history placed in its social context is, in principle, a desirable pursuit. But I admit that I vastly prefer intellectual history on its own to social history adrift from intellectual moorings. When the Agassiz Museum was in its early prime, this social history serves as a gloss on intellectual history generally. Not so, though, after the overall quality of the intellectual enterprise had declined. It is a strength of Winsor's approach that she lets us see this, but, as a satisfying read, I still prefer the first half of the book.

Above all, Winsor shows that the strength of the scientific enterprise of systematics at an institution reflects far more its staff's collective ability and connectedness with the outside world than any intrinsic aspect of the sociopolitical structure of the institution itself. Occasionally, bursts of creative insight do emanate from centres that house a talented group of systematists, but only when the problems they address are central to the discipline as a whole and their new ideas are sufficiently well

crafted to attract the attention of the entire systematics community. There is no substitute for talent — a message that all who are concerned with the health of today's natural history museums should take to heart. □

Niles Eldredge is in the Department of Invertebrates, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024-5192, USA.

■ Two recent books examine the 'coming of age' of American biology. *The Expansion of American Biology* (eds K. R. Benson, J. Maienschein and R. Rainger) covers critical developments between the two world wars (Rutgers University Press, \$42 (hbk), \$17 (pbk)); *Transforming Traditions in American Biology, 1880–1915* by J. Maienschein focuses on the work of E. B. Wilson, E. G. Conklin, T. H. Morgan and R. S. Harrison (Johns Hopkins University Press, £34.50).



Field sketch by Jacques Burkhardt (1865) of the Amazonian *Pterophyllum scalare*, the 'angelfish' of later aquarium hobbyists. The drawing is annotated by Louis Agassiz. (Courtesy of the Museum of Comparative Zoology, Harvard University.)

talent. Winsor admirably presents a parable on echinoderm systematics to parallel the crayfish story. We read of the "bombshell" that Alexander dropped at a meeting of the American Association for the Advancement of Science: a probability argument, based on his struggles with sea-urchin systematics, that phylogenetic pattern is in principle unrecoverable. Suffice it to say, Alexander's bombshell was no lasting contribution to systematic biology.

Winsor is riveting when dwelling on scientists locked in genuine intellectual battle with the world at large. Unfortunately, when dealing with lesser lights, she is equally at the mercy of the characters. Minor players, peripheral to the intellectual mainstream, simply cannot be as arresting or serve as equally instructive examples of 'how science is

Towering infernos?

Jonathan Fink

Mountains of Fire: The Nature of Volcanoes. By Robert W. Decker and Barbara B. Decker. Cambridge University Press: 1991. Pp. 198. £30, \$39.95 (hbk); £10.95, \$15.95 (pbk).

VOLCANOES receive some of the world's worst publicity. Journalists and geologists focus on lurid details of each new disaster; the former to sell newspapers, the latter to attract funding for hazard assessments. Yet, as Robert and Barbara Decker convincingly point out, the study of volcanoes extends far beyond the tallying of deadly statistics. In this comprehensive and enjoyable book, we learn to view these dynamic "fire mountains" as instruments for bringing about climate change, mineral wealth, agricultural renewal and even biological evolution. At the same time, the dangers of eruptions are thoroughly documented and explained.

The past two decades have witnessed a blossoming of understanding about volcanic processes owing mainly to long-term monitoring of Kilauea and Mauna Loa in Hawaii, Mount St Helens in Washington State and Mount Etna in Italy. Other important advances have come from the study of spacecraft images of Mars, the outer planet satellites and Venus, as well as from observations of the mid-ocean ridges. The Deckers have been involved with many of these findings, he as scientist-in-charge at the United States Geological Survey's Hawaiian and Cascades Volcano Observatories, and she as the coauthor of numerous books and articles that covered these events. Their experience is reflected in the breadth and timeliness of the topics they deal with in this book.

The reader is drawn into each chapter by an eyewitness account of a significant eruption or volcanological discovery. These anecdotes include familiar descriptions of the destruction of Pompeii in AD 79 and the explosion of Mount St Helens in 1980. Less well known but equally exciting are the first glimpses of active volcanism on Jupiter's moon Io, as viewed by NASA scientists at the Jet Propulsion Laboratory in 1979, and the bizarre menagerie of creatures surrounding 'black smoker' hot springs on the East Pacific Rise, seen through the windows of a small research submarine in 1977. Vivid narrations also convey the tragic shortsightedness of politicians and bureaucrats, which greatly increased the fatalities associated with the two deadliest eruptions this century, those of

Mont Pelée in 1902 and Nevado del Ruiz in 1985.

A considerable challenge for any popular book about geology is to translate the spatial and temporal scales of the Earth's dimensions and history into the reader's frame of reference. The Deckers are uncommonly effective at using analogies to make these connections. Continents drift apart about as fast as fingernails grow, and the Washington Monument could fit easily inside the Mount St Helens lava dome. Every page is richly embellished by such images.

Simple maps, cartoons and cross-sectional diagrams help the authors to minimize technical jargon. Sixteen pages of colour plates strikingly portray many common eruption phenomena. Black-and-white photographs, liberally scattered throughout the book, are less consistent in quality. Many appear to have been made from colour slides, leading to low contrast and rather poor focus.

Shifting populations

John Brookfield

Molecular Genetic Ecology: In Focus. By A. Rus Hoelzel and Gabriel A. Dover. Oxford University Press: 1992. Pp. 100. £6.50, \$14.95 (pbk).

I HAD hoped that I would be able to recommend this book, not least because of the increasingly large community of ecologists who are starting to use molecular genetic techniques to investigate population structure. These workers require a short text that will tell them what they need to know about molecular and population genetics. But I do not think this book is adequate for their needs.

The areas dealt with include the structure of eukaryotic genomes and the techniques for their study. There is a summary of theories of molecular evolution, the statistical interpretation of genetic surveys and some examples of the uses of molecular techniques to study animal populations. The sections on molecular biology and the final glossary are quite good, although permeated by confusing references to molecular drive. The authors seem unclear about the extent to which they think inheritance of DNA is biased. They consistently warn that drive processes may be spreading sequences through populations, yet simultaneously advocate analytical approaches that implicitly assume that substitutions in evolution are neutral, such as the estimation of mutation rates from substitution rates.

The authors deal in some detail, considering the book's length, with evolu-

Nonetheless, the broad range of features that they show adds another important dimension to the text.

One of the ironies of volcanic research is that although our ability to predict eruptions continues to increase, the number of casualties does not appreciably decline. This is largely the result of growing populations around tropical volcanoes, drawn to soils enriched by the breakdown of ash and lavas. But many deaths, including nearly all of the 25,000 in Armero, Colombia, could have been avoided with simple precautions if civic leaders had grasped the dangers. *Mountains of Fire* will go a long way towards educating these people. It can also help those of us who hear about these disasters in the news to appreciate more fully their causes and consequences. □

Jonathan Fink is in the Department of Geology, Arizona State University, Tempe, Arizona 85287-1404, USA.

tionary theory and population genetics, and consider the statistical interpretation of molecular data. These sections are the weakest parts, and contain many errors. Two examples will suffice. In describing how one arrives at Nei's distance, D , the authors say that a "simple measure of genetic similarity (I) is multiplied by the natural logarithm ($\ln$) to give a parameter that is 0.0 for genotypes that are completely dissimilar." Later, after a description of the variances of heterozygosity between and within loci, they add the qualification "[these] variance measures are affected by . . . dominance. Dominance will tend to increase the variance, although the effect is small unless the frequency of recessive genes is very small." Reading passages like these, one inevitably wonders why the authors have chosen to write a book including subjects that they know so little about.

There is scope for a better book on this subject, one that will not only be more accurate but also pay more attention to the question of the statistical power of small samples. Large surveys on various animal populations using DNA markers frequently turn out to have disappointing and equivocal results, largely because they are carried out thoughtlessly by field workers who are not aware of the large sample sizes needed to establish general principles of population structure and behaviour. □

John Brookfield is in the Department of Genetics, University of Nottingham, Queens Medical Centre, Nottingham NG7 2UH, UK.

■ Columbia University Press has published the second edition of *The Evolutionary Process: A Critical Study of Evolutionary Theory* by Verne Grant. For a review, see *Nature* 320, 317; 1986. Price \$60. The book now covers molecular evolution.

NATURE · VOL 355 · 20 FEBRUARY 1992

Nature's jewels

Colin W. Harding

The Daguerreotype: Nineteenth-Century Technology and Modern Science. M. Susan Barger and William B. White. *Smithsonian Institution Press: 1991.* Pp. 252. £39.95, \$67.95.

SIR John Herschel, a man not given to hyperbole, was almost unable to contain his enthusiasm when he first saw them: "It is hardly saying too much to call them miraculous. Certainly they surpass what I could have conceived as within the bounds of reasonable expectation." The 'miracles' that Herschel had witnessed were some of the earliest examples of daguerreotypes, the first practical method of making photographs.

The daguerreotype process was announced to the world at the joint meeting of the Académie des Sciences and the Académie des Arts in Paris on 19 August 1839. Invented and named with a showman's modesty by the French artist Louis Jacques Mandé Daguerre, daguerreotypes were produced on silver-plated copper sheets sensitized with iodine vapour and 'developed' in mercury vapour, producing jewel-like images that possessed a verisimilitude that contemporary observers viewed with disbelief. Samuel Morse remarked: "The exquisite minuteness of the delineation cannot be conceived. No painting or engraving ever approached it." As an image produced solely by the agency of the forces of nature, the daguerreotype was seen not as a representation of nature but as nature itself.

Although the daguerreotype found its greatest application in commercial portraiture, it was clear from the beginning that the technique also had great potential as a scientific tool, offering a means of objective, accurate and detailed recording. The chemist Joseph Gay-Lussac realized as early as 1839 that "it is certain that through Monsieur Daguerre's invention physics is today in possession of a reagent extraordinarily sensitive to the influence of light, a new instrument which will . . . furnish the nucleus around which new researches and new discoveries will be made."

The daguerreotype was to be superseded by other photographic processes by the 1850s, but for a while it did indeed prove to be a valuable tool in many areas of study. The question of how daguerreotype images were formed, however, remained unresolved by nineteenth-century scientists, for the technology of the daguerreotype was far in advance of the science needed to describe how the process worked and the image was formed.

Barger and White have now produced a fascinating and comprehensive book on this early photographic process from the perspective of materials scientists. Drawing on the results of more than six years of work using the latest laboratory techniques, they have succeeded in unlocking the mystery of the daguerreotype, revealing the complex interplay of chemistry and optical physics that lies behind the creation of the sublimely beautiful images. As a mixture of historical research, scientific investigation and practical advice, this is, in effect, three books in one.

The authors start with the history of the daguerreotype, tracing its invention and development. They examine how the process was used by nineteenth-century photographers and scientists and detail early efforts to understand and



Daguerreotype of Louis Daguerre (1844).

explain the process. In the central section, the authors summarize the results of their own work into the physics and chemistry of how daguerreotype images are formed, showing that they arise from a characteristic microstructure of small silver spheres dispersed over a polished silver surface. In the final chapters, aimed primarily at conservators and curators, the authors describe both old and new methods of cleaning, preserving and displaying daguerreotypes.

Barger and White address the book to "historians of technology and science, to the curious general reader, and most important, to those charged with the care and preservation of photographs." Such is the scope of the authors' work that each of these groups will find much in the book that is useful. It is a valuable addition to photographic historiography and should stimulate similar research on other early photographic processes. Although the secrets of the daguerreotype have now been revealed, the beauty and fascination of the images remain undiminished. □

Colin W. Harding is curator of the Kodak Museum at the National Museum of Photography, Film and Television, Prince's View, Bradford BD5 0TR, UK.

New in paperback

- *When Did I Begin?* by Norman Ford examines the concept of the human individual in history, philosophy and science. Published by Cambridge University Press, price £9.95, \$16.95.
- *Between Two Worlds: Science, the Environmental Movement and Policy Choice* by Lynton Keith Caldwell. Published by Cambridge University Press, price £12.95, \$19.95.
- Dover have republished *The Magic of Numbers* by Eric Temple Bell. First published in 1946, the book surveys the origins and influence of mathematical thought. Price £8.95, \$8.95.
- *Palaeobiology: A Synthesis* edited by Derek E. G. Briggs and Peter R. Crowther contains more than 120 review articles. Invaluable reading for all researchers and students in the field. Published by Blackwell Scientific, price £34.50.
- Penguin have published their *Dictionary of Information Technology and Computer Science* by Tony Gunton. Price £5.99.
- *In To Save an Elephant*, Allan Thornton and Dave Currey of the Environmental Investigation Agency tell how they infiltrated and exposed the illegal ivory trade. Published by Bantam, price £5.99.

New editions

- *Principles of Biochemical Toxicology* by J. A. Timbrell (2nd edn). Published by Taylor and Francis, price £19.95 (pbk). For a review, see *Nature* **302**, 179; 1983.
- *Understanding DNA and Gene Cloning: A Guide for the Curious* by Karl Drlica (2nd edn) has introductions by Bruce Alberts and Arthur Kornberg. Published by Wiley, price £17.95 (pbk).
- *Nerve and Muscle* by R. D. Keynes and D. J. Aidley (2nd edn) is a concise and readable introductory textbook. Published by Cambridge University Press, price £27.95, \$49.95 (hbk), £9.95, \$15.95 (pbk).
- *Microbes and Man* by John Postgate (3rd edn) has, since publication of the first edition in 1969, become widely acclaimed as a classic. Written for the general reader, the book is also a superb introduction for students of microbiology. Published by Cambridge University Press, price £24.95, \$49.95 (hbk), £7.95, \$15.95 (pbk).
- *Sedimentary Petrology* by Harvey Platt (2nd edn). Published by Freeman, price £38.95. For a review, see *Nature* **302**, 183; 1983.
- *Thinking, Problem Solving, Cognition* by Richard E. Mayer (2nd edn). Published by Freeman, £29.95 (hbk), £15.95 (pbk). For a review, see *Nature* **308**, 131; 1984.
- *Volume I of High Energy Astrophysics* by Malcolm S. Longair (2nd edn) provides a comprehensive introduction to particles, photons and their detection. Published by Cambridge University Press, price £45, \$69.95 (hbk), £16.95, \$34.95 (pbk). For a review, see *Nature* **295**, 461; 1982.
- *Handbook of Toxic and Hazardous Chemicals and Carcinogens* by Marshall Sittig (3rd edn) is published by Noyes Data Corporation. 2 Volumes, price \$197.
- *Modern Spectroscopy* by J. M. Hollas (2nd edn) is aimed at students of physics and chemistry. Published by Wiley, price £14.50 (pbk), £39.50 (hbk).

Memoirs of a playful fox

Ursula Mittwoch

Odyssey of a Scientist: An Autobiography. By Hans Kalmus. Weidenfeld and Nicholson: 1991. Pp. 166. £17.99.

HANS Kalmus died in November 1988, only a few months after he had begun a collaboration for this 'autobiography' with the author and broadcaster David Benedictus. Together they had completed two chapters, the remaining nine lying in a vast mass of files, notebooks and diaries. Benedictus tells us that he wishes the book to read like Hans Kalmus and not like David Benedictus, and in this he has undoubtedly succeeded.

Kalmus was born in Prague and studied zoology and medicine at the German University where he became a teacher. In 1939, at the age of 33, he arrived in England, shortly before the German army invaded Prague, and soon obtained a research post at University College London. Because his membership of University College lasted close to 50 years, during most of which he lived in Harpenden, a small town some 20 miles north of London, the "Odyssey" of the title may be regarded as something of a hyperbole. He did on occasion travel off the beaten track, but on the whole his life was enacted in a more minor key.

The Harpenden connection was a result of the Second World War, when University College was evacuated — a wise precaution as it turned out, as the college suffered a direct hit in September 1940, and a second bomb attack in April 1941. J. B. S. Haldane, who had offered Kalmus a place in his department, at first refused to move, but had to give in when water, electricity and telephones were disconnected. Research facilities were negotiated at Rothamsted Experimental Station and a house was rented nearby. The occupants included Haldane, Helen Spurway, who was a former student and his future second wife, his secretary and some other members of the department. Into this commune moved Hans Kalmus, his wife Nussy (pronounced 'Nushy'), two small sons and a grand piano. Haldane did not enjoy Hans's playing, preferring to entertain the residents with readings of Byron's *Don Juan* and *Childe Harold*.

During working hours, Haldane led a

civilian team investigating escape from submarines, a team that Kalmus was invited to join. The experiments involved subjecting themselves and others to different gas mixtures at pressures up to 10 atmospheres. This was hazardous, and there were several mishaps, though none of them fatal.

Kalmus was a prolific writer; before moving to England he had already published a book on *Paramecium*. One night, as he was fire-watching with Lettice Crump, a protozoologist from Rothamsted, she remarked to him that nobody could understand genetics; and so they decided to do something about it. On subsequent fire-watching evenings, and without consulting a book,

as a new graduate in 1947, I absorbed the impression of the time that writing science for the general public was definitely *infra dig*, a kind of prostitution. True, Haldane was an acknowledged popularizer, but his position placed him outside this particular criticism.

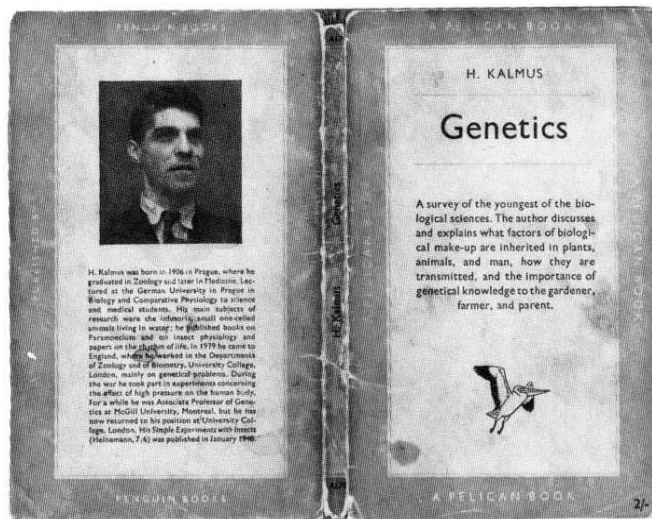
Furthermore, in contrast to Haldane, Kalmus was not a natural popularizer. His command of English remained patchy, and he was not interested in moulding words and sentences into clear and concise statements. As he himself said, the success of *Genetics* was partly due to the fact that its publication coincided with news of the Lysenko affair, a manifestation of the state-dominated genetics in stalinist Russia. In addition, at one shilling and sixpence, this paperback was a bargain for schools and individuals with a serious interest in genetics. As for the general public, then as now, it seemed destined to buy rather than read scientific bestsellers.

Eventually Kalmus became a professor of biology. The title reflects the catholic nature of his research, which is also demonstrated by his published books: *Paramecium*; *Genetics*; *Simple Experiments With Insects*; *Variation and Heredity*; *The Chemical Senses in Health and Disease*; *Biological Rhythms*; *Diagnosis and Genetics of Colour Vision*; *Regulation and Control in Living Systems*; and, finally, a volume of poetry in German.

Kalmus recalls that Karl von Frisch, famous for his discovery of dancing bees, once said of him: "If only this young man could stick to one subject." Kalmus never managed to do so, maintaining that the characteristic to which he owed whatever success he had was his playfulness.

Another emigré scholar, Isaiah Berlin, has classified intellectuals into two groups, hedgehogs and foxes, on the basis of a Greek fragment, "The fox knows many a thing, but the hedgehog knows one big thing." There can be no doubt that Kalmus was a fox, pursuing many ends, often unrelated and even contradictory. In present-day biology, this breed is doomed to extinction, their habitat being taken by the subspecies of molecular hedgehogs. This, surely, gives added interest to these memoirs, not only for those who knew Kalmus, but also to others with a serious interest in scientific research. □

Ursula Mittwoch is in the Department of Anatomy, Faculty of Basic Medical Sciences, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK.



Genetics for the masses — an early reprint of Kalmus's bestseller.

Hans dictated what he thought to be the basic facts necessary for an understanding of genetics, while Lettice wrote it all down and corrected the English. The manuscript was published as a Pelican book in 1948 and became a bestseller.

The success of *Genetics* gave Kalmus a taste of celebrity life. People recognized him from the photograph on the book's cover, and successful junior colleagues told him that they had become interested in genetics through reading his book. "But the popularity of a scientist", he writes, "cannot compare . . . with that of footballers or actors. As proof I can attest that I have been far more frequently mistaken for Burt Lancaster, whom I must somehow resemble, than recognized for myself."

But the book did nothing to improve his position at University College. After the war, Lionel Penrose became Galton Professor and head of the Galton Laboratory, of which Kalmus became a member. When Kalmus received an offer of an associate professorship at McGill University, Penrose advised him to accept. Kalmus was reluctant, but went to McGill for a year.

Having joined the Galton Laboratory

Supernova 1987A at five years of age

Roger A. Chevalier

Supernova 1987A is gradually changing from a supernova heated by the radioactive decay of heavy elements to a supernova remnant interacting with a surrounding medium whose structure has already been revealed through its illumination by radiation from the explosion. The nature of the compact remnant remains a central mystery.

THE last known supernova in our own Galaxy was Cassiopeia A, which exploded in about 1680. Other galactic supernovae have probably occurred since then, but have been invisible to us because of obscuration by dust. Although supernova 1987A was not in our Galaxy, its location in a nearby unobscured companion galaxy, the Large Magellanic Cloud (LMC), has made it the first local supernova to have been studied with all the resources of modern observational astronomy. SN1987A occurred not only at a propitious location, but also at a propitious time: technologies of quite recent origin allowed new types of supernova observation, notably neutrino detection, γ -ray spectroscopy and infrared spectroscopy. This range of observational methods has produced a remarkably complete description of the event. Here I review the observations and our understanding of the supernova event and its aftermath, dividing the subject into the different types of phenomena that were observed, and emphasizing recent results.

The progenitor star

From comparison of position of the supernova with previous photographs of the field, the progenitor was identified as Sk - 69 202 in the Sanduleak list of Magellanic Cloud stars. The current bolometric luminosity of the supernova, $1 \times 10^{37} \text{ erg s}^{-1}$ (ref. 1), is below the bolometric luminosity of Sk - 69 202, $\sim 4 \times 10^{38} \text{ erg s}^{-1}$, proving that the star did explode. The progenitor was a star of spectral type B3 and luminosity class I (ref. 2), a blue supergiant, with a radius of $\sim 3 \times 10^{12} \text{ cm}$. In light-curve models, this initial stellar radius gives a good fit to the observed light curve^{3,4}. The required explosion energy is $(1.4 \pm 0.6) \times 10^{51} \text{ erg}$.

The blue progenitor was a surprise. Massive stars ($\geq 10 M_{\odot}$) had been expected to explode as extended red supergiants if they retained their hydrogen envelopes. Models of exploding red supergiants agreed well with light curves of observed type II supernovae. SN1987A showed hydrogen lines in its spectrum, and the light curve implied a massive hydrogen envelope, as expected for a type II supernova. Its explosion as a blue star stimulated new work in massive star evolution. An additional clue was the detection of dense gas near the supernova by its line emission and dust scattering^{5,6}. The high density suggested mass lost during a red supergiant phase which ended $\sim 10^4$ years before the explosion. With a particular choice for the treatment of convection (restricted semiconvection) and heavy element abundances appropriate to the LMC (1/3 of solar abundances), stellar evolution models for a star of initial mass $20 M_{\odot}$ have been able to produce the expected evolution to the red supergiant phase, with a move to the blue $\sim 10^4$ years before the explosion^{3,7,8}. Such models do not accurately reproduce the observed distribution of stars in the luminosity-temperature diagram for the LMC, especially the number of blue supergiants close to the main sequence, but other factors, such as binary evolution, may play a part. Although these models are sensitive to input parameters, they do produce explosions as a blue supergiant in a natural way for LMC abundances.

The circumstellar matter around SN1987A has a large overabundance of nitrogen compared with cosmic material⁵, and helium may be enhanced in the circumstellar⁹ and supernova¹⁰

gas. These abundance anomalies are not naturally produced in the above models. Saio *et al.*¹¹ have proposed models with unrestricted convection, substantial mass loss, and mixing of He and N from the core into the envelope by an *ad hoc* mechanism. Mixing by meridional circulation is possible, but would require the star to be rotating rapidly, which is unlikely while it is in an evolutionary state with an extended radius. It is not clear whether this mechanism will produce supernovae from red supergiant stars for solar abundances.

An alternative view is that binary evolution has been important. Mass transfer onto a given massive core can force evolution to the blue^{12,13}. If close binary evolution has occurred (probably in a common envelope phase), angular momentum can be transferred to the stellar progenitor envelope while it is in an extended phase¹⁴, possibly helping to mix the envelope as well as to produce the observed asymmetries in the ejecta and in the circumstellar medium (see below). But direct evidence for a binary companion is lacking.

The central compact object

The neutrino burst detected from SN1987A^{15,16} on 23 February 1987 confirmed the basic theory of core collapse and refined our views of neutrino properties (see ref. 17 for a review). Twenty neutrinos in the energy range 6 to 39 MeV were detected over $\sim 12 \text{ s}$. The timescale for the emission was consistent with predictions allowing for neutrino trapping in the collapsed core. The total flux of neutrinos was consistent with neutrinos carrying off the binding energy of a central neutron star, and the neutrino energies, after allowing for the detector response, were consistent with thermal emission from a region the size of a neutron star.

The duration of the event showed that a neutron star, rather

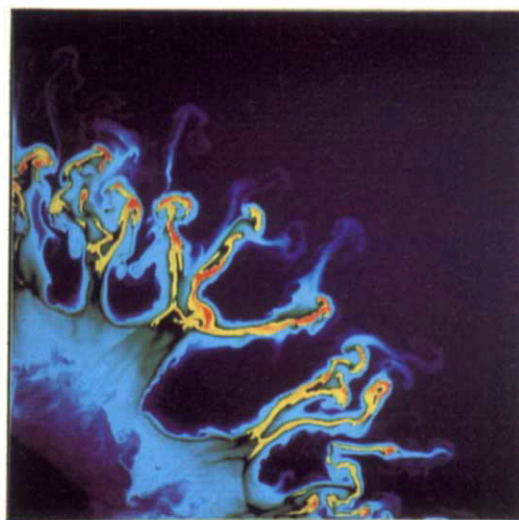


FIG. 1 The oxygen abundance structure of a two-dimensional calculation of SN1987A at an age of 3.5 hours. The complex structure results from hydrodynamic instabilities at the composition interfaces (from ref. 36, courtesy of W. D. Arnett).

than a black hole, formed during the initial collapse phase¹⁸. But some initially expanding gas is likely to be bound to the neutron star and to fall back to the surface, where strong neutrino losses allow it to be added to the neutron star¹⁹. In a spherically symmetric model, $\sim 0.1 M_{\odot}$ of matter might accrete to the neutron star in hours after the explosion²⁰. Depending on the initial mass of the neutron star and the equation of state for the dense matter, this may cause collapse to a black hole.

In the spherical accretion model, radiation is initially trapped in the dense accretion flow and only a small luminosity ($\sim 3 \times 10^{34} \text{ erg s}^{-1}$ after a year) emerges from outside the radiation trapping radius. If a black hole has formed, the luminosity slowly drops in subsequent years. If a neutron star is present, the shocked region eventually extends to where the shock front reaches the trapping radius at an age of about one year, and the escaping luminosity rises to the Eddington limit, $\sim 3 \times 10^{38} \text{ erg s}^{-1}$. The neutron star envelope then loses energy by radiation and is radiation-supported. Mass loss may occur during this phase. The upper limit on a steady luminosity source in SN1987A is now below $10^{37} \text{ erg s}^{-1}$ (ref. 1), showing that such a radiation-supported accretion envelope is not present. Either a black hole has formed or the accretion envelope has dissipated.

The possibility that angular momentum of accreting material might give rise to a disk around the neutron star is of interest both for pulsar theory²¹ and for planet formation around neutron stars²². In this case, radiation should be able to escape more easily, and the absence of an Eddington flux is perplexing unless the disk has rapidly evolved to a state where very little accretion occurs. We know little about young neutron stars and cannot be certain that they have strong magnetic fields in the first few years after their birth²³. If a strong magnetic field is present (10^{12} gauss is typical), the electromagnetic spin-down power of the pulsar can be expected to create a cavity around the neutron star once the matter and radiation pressures become smaller than the magnetic pressure. The pulsar bubble will eventually sweep up mass that is part of the uniformly expanding supernova gas and drive a shock front into the supernova. The radiative shock front emits $\sim 0.015 \dot{E}$ (ref. 24), where $\dot{E}$ is the spin-down power of the pulsar. If this is all that is emitted, the current limit of $10^{37} \text{ erg s}^{-1}$ on a constant energy source in SN1987A implies a limit of $7 \times 10^{38} \text{ erg s}^{-1}$ on $\dot{E}$. Observations of young pulsar nebulae show, however, that $\sim 10\%$ of $\dot{E}$ can be emitted by the synchrotron X-ray nebula around the pulsar²⁵. If this result from 1,000-year-old pulsar nebulae can be applied to those only a few years old, the upper limit on $\dot{E}$ is reduced to $10^{38} \text{ erg s}^{-1}$.

If a neutron star is present, a minimum luminosity is the black-body luminosity from the neutron star surface. At an age of 5 years, the expected temperature is $\sim 2.5 \times 10^6 \text{ K}$ (ref. 26) and the luminosity is $3 \times 10^{34} \text{ erg s}^{-1}$. Most of the emission is in the soft X-ray range, which is heavily obscured by the expanding supernova gas. Optical depth unity is expected at an age of ~ 300 years.

The supernova ejecta

Models for the early light curve of SN1987A showed excellent agreement with expectations for the explosion of a blue supergiant star. After day 40, some other energy source was needed for the light³, and it became clear after day 120, when the light curve entered an exponential decay with a half-life of 77 days, that the energy source was the radioactive isotope ^{56}Co . This was not unexpected because both theory²⁷ and observation²⁸ suggested that previous type II supernovae were powered by ^{56}Co decays at late times. A new possibility here was the direct detection of the high-energy radiation responsible for the light curve. The power is initially in the form of γ -rays with an energy of $\sim 1 \text{ MeV}$, which are scattered by electrons in the expanding gas and downgraded in energy. The electrons deposit their energy in the thermal gas; it is assumed that this occurs locally over the timescales of interest. The photons scatter down in

energy until photoabsorption becomes the dominant interaction process and the photon energy is thermalized. The energy at which this transition occurs depends on the composition of the gas and occurs at $\sim 20 \text{ keV}$ for a solar composition²⁹. Initially the photon energy is trapped in the gas, and the radioactive energy input is radiated as optical light from the photosphere. Eventually, as the photons can diffuse further out in the expanding gas, the downscattered photons are able to escape in the X-ray range. The X-ray spectrum is expected to have an energy cut-off where the photons begin to interact primarily by photoabsorption. Finally, the γ -rays escape directly and gradually evolve to their full input intensity.

Observations of SN1987A were generally in line with these expectations. X-ray observations showed a high-energy spectrum with a low-energy cut-off at $\sim 20 \text{ keV}$ (refs 30, 31). The ^{56}Co lines at 0.843 and 1.238 MeV were first observed by the Solar Maximum Mission³². The problem was that both the X-rays and γ -rays turned on at 6 months, several months earlier than had been predicted by spherically symmetric models in which all of the ^{56}Ni , the progenitor nuclide to ^{56}Co , ends up at the very centre of the explosion. The high-energy photons should thus take a long time to diffuse out and be observed. The early escape of the high-energy radiation implied that the heavy elements were mixed into the supernova envelope gas. Model fits to the optical light curve were also improved by allowing mixing of the radioactive energy source^{3,4}. The modelling of the optical and high-energy light curves included mixing in an *ad hoc* way and retained spherical symmetry. The fact that the X-ray spectrum extended to $\sim 20 \text{ keV}$ showed that the mixing was not uniform because the enhanced heavy element abundance would give a higher energy cut-off. The ^{56}Co must be in clumps from which the photons propagate into hydrogen-rich gas. Clumping is also indicated by analysis³³ of the [O I] lines at 6,300 and 6,363 Å. Such a situation is shown by two-dimensional hydrodynamic simulations of the explosion³⁴⁻³⁶ (Fig. 1). The models of ref. 34 assume that the initial explosion is spherically symmetric, but then change to a two-dimensional grid while the shock front is still propagating through the star. Hydrodynamic instabilities related to the Rayleigh-Taylor instability occur in regions where there is a steep drop in density

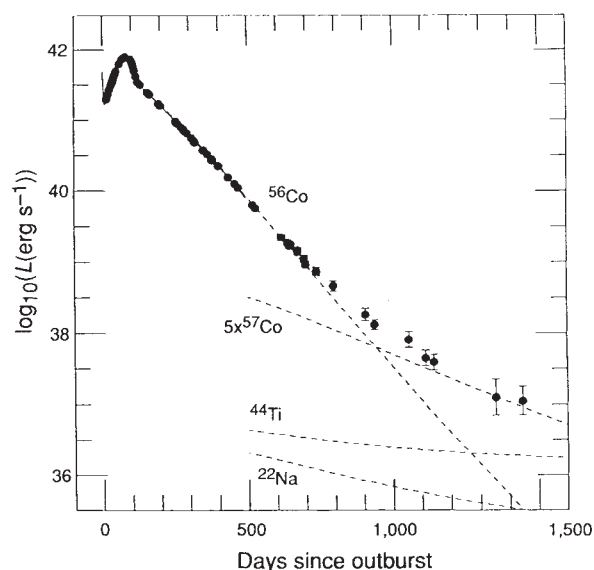


FIG. 2 The luminosity of SN1987A, from ultraviolet to infrared, measured at the Cerro Tololo Inter-American Observatory through day 1,444. The energy deposition by the radioactive nuclides (dashed lines) is based on the following initial masses: ^{56}Ni (and subsequently ^{56}Co), $0.075 M_{\odot}$; ^{44}Ti , $1 \times 10^{-4} M_{\odot}$; ^{22}Na , $2 \times 10^{-6} M_{\odot}$; and ^{57}Co , $0.009 M_{\odot}$ ($5 \times$ the value expected from the solar ratio of $^{57}\text{Fe}/^{56}\text{Fe}$) (from ref. 1, courtesy of M. Phillips).

because relatively dense gas is decelerated by the surrounding lower-density gas. The steep density gradients occur in regions of composition discontinuities, both the He/H and the O/He discontinuities. The complex structures seen in Fig. 1 are the result of dense Rayleigh–Taylor fingers of gas being disrupted by Kelvin–Helmholtz instabilities in the velocity shear layers at the finger edges.

Two-dimensional calculations run for the first few hours of supernova time, during which a free expansion phase begins, give a maximum velocity for the radioactive species³⁷ of $v_{\max} \approx 1,300 \text{ km s}^{-1}$, which can be compared with the observed $v_{\max} \approx 3,000 \text{ km s}^{-1}$ in the ^{56}Co lines^{38,39}. Allowing for three-dimensional motion increases the result by $\leq 20\%$. A possible additional acceleration mechanism is the energy input from ^{56}Ni , which decays with a half-life of 6 days. Herant and Benz³⁷ have carried instability calculations to later times and include this mechanism, but find that the velocities increase by only $\sim 30\%$ and remain smaller than the observed values. A possible solution is mixing associated with the initial explosion. If type II supernova explosions are initiated by a neutrino-heated bubble around the proto-neutron star, the interface between the hot bubble and the surrounding gas is Rayleigh–Taylor unstable so that mixing is expected⁴⁰. Numerical calculations⁴¹ have borne out this expectation for evolution on a timescale of seconds. The mechanism has not yet been combined with the mechanisms that occur at later times. Multidimensional models are needed for accurate calculation of light curves from optical to γ -ray wavelengths for SN1987A, as well as the line profiles. The inadequacy of the one-dimensional mixed models is shown by the fact that their predicted widths for the Fe group lines⁴² are $\sim \frac{1}{2}$ those observed^{38,39,43}.

Although these instabilities have a characteristic length scale that is a fraction of the stellar radius, there is evidence for large-scale asymmetry in SN1987A. Profiles of the γ -ray ^{56}Co lines show a systematic redshift relative to slow circumstellar gas of $\sim 500 \text{ km s}^{-1}$ (refs 38, 39); there is a global asymmetry in the ^{56}Co distribution. The observed intrinsic polarization in the optical light⁴⁴ implies an asymmetry in the expanding envelope of $\sim 10\%$ (ref. 45). A similar magnitude of asymmetry is needed in either the initial envelope or the core explosion¹⁴. A possible reason is rapid rotation of the stellar envelope, which, in turn, probably requires a binary companion so that angular momentum can be transferred to the envelope at a late stage of evolution. Attractive features of this hypothesis are meridional mixing in the presupernova star and a possible explanation for asymmetric circumstellar gas. Alternatively, a core asymmetry, already indicated by the asymmetry in the ^{56}Co emission line, can propagate through the envelope without attenuation during the explosion^{14,46}. The fact that galactic pulsars, thought to originate in type II supernovae, have typical velocities of $100\text{--}200 \text{ km s}^{-1}$ gives an additional reason to expect global asymmetries in the initial explosion.

Ejecta emission properties

The brightness of the explosion yielded an abundant harvest of spectra at ultraviolet to infrared wavelengths and thus stimulated detailed work on supernova atmospheres. Non-LTE (local thermodynamic equilibrium) effects were included and found to be important particularly for H and He (ref. 10). Eastman and Kirshner¹⁰ modelled the spectrum for the first 10 days and derived a distance of $49 \pm 6 \text{ kpc}$ to the LMC, in good agreement with distances derived from more standard methods. Application of the same method to more distant supernovae is a promising way of establishing the distance scale in the Universe.

Four months after the explosion, SN1987A entered the nebular emission phase, in which line emission dominates continuous emission in the spectrum. The spectrum is produced by the deposition of energy from the energetic photons resulting from radioactivity⁴⁷. Modelling the emission spectrum involves calculating the energy deposition throughout the supernova and

solving for thermal and ionization balance at each point. Energy deposition by high-energy photons tends to give low ionization over a broad region. From such calculations, Fransson and Chevalier⁴⁸ predicted the line emission from the heavy element supernova core for the first few years, using a massive star explosion model of Woosley. The predictions for the strengths and evolution of the [O I] and [Ca II] lines fit the observations over the first 500 days quite well^{49,50}. Subsequently, the models predicted a sharp drop in line flux because of a shift in cooling from transitions between multiplets to transitions between fine-structure levels. The cooling becomes insensitive to temperature in the range of a few 100 to a few 1,000 K, so that the temperature drops by a large factor when there is a small drop in the heating rate. Because the emission shifts to infrared fine-structure lines, this transition is referred to as the ‘infrared catastrophe’⁹⁸. Observations of SN1987A showed an initial drop of the optical lines near day 500, but the fluxes subsequently remained higher than predicted. At about the same time, there was a widening of the optical lines⁵¹, suggesting emission from gas further out in supernova. It may be that core material did cool at this time, but envelope material, with its lower emissivity, remained warm and came to dominate the spectrum.

There is evidence that some gas did cool to low temperatures at an age of 500 to 600 days. From an analysis of near-infrared lines, Spyromilio and Graham⁵² estimated the mass of Fe in the ejecta to be less than the $0.075 M_{\odot}$ of ^{56}Fe expected from decays of ^{56}Co , and they suggested that the missing Fe cooled to the point where it no longer emitted near-infrared lines. Strong cooling is also implied by the evidence for dust formation at an age of 500–600 days. At this time, there was a drop in the optical flux and a rise in the infrared flux⁵³ so that the total luminosity continued to track that expected from ^{56}Co decays. At the same time, the emission lines showed absorption of redshifted gas, implying that there was some absorbing medium mixed in with the emission⁵⁴. The most plausible explanation for these observations is dust formation, which requires cool temperatures. To explain the wavelength dependence of the shifts in emission line peaks, Lucy *et al.*⁵⁴ suggest the presence of a diffuse dust component and an optically thick component with a small filling factor. The required mass of emitting dust is small, $\sim 10^{-4} M_{\odot}$. Kozasa *et al.*⁵⁵ have shown how the dust might nucleate in the expanding supernova.

Observations of the shift to the infrared are difficult because of atmospheric effects and because the detector technology is only now becoming available. This has resulted in uncertainties in the late-time luminosity curve for SN1987A. At an age of $\sim 1,000$ days, results from the European Southern Observatory (ESO)⁵⁷ gave a luminosity about a factor 2 larger than those from the Cerro Tololo Inter-American Observatory (CTIO)⁵⁷ because of differences in the 10- and 20- μm fluxes. Figure 2 shows the luminosity curve measured from CTIO to an age of 1,444 days. Beginning at an age of 900 days, there was a flattening of the light curve from the ^{56}Co decay curve. Such a flattening had been predicted on the basis of the power provided by decays of ^{57}Co , which has a half-life of 271 days⁵⁸. Nucleosynthesis calculations suggested that the $^{57}\text{Co}/^{56}\text{Co}$ ratio should correspond to 1.5–2.5 times the solar ratio of $^{57}\text{Fe}/^{56}\text{Fe}$ (refs 59, 60). As shown in Fig. 2, the late-time decay rate is well fitted by the ^{57}Co decay rate but with a $^{57}\text{Co}/^{56}\text{Co}$ ratio 5 times the solar value. Suntzeff *et al.*¹ find that subtracting the fitted ^{56}Co decay curve at early times from the total luminosity curve results in a luminosity profile with the ^{57}Co decay rate. This persists over a decrease in luminosity by a factor of 30. Although this seems to be excellent evidence for ^{57}Co decay, there may be observational problems as well as theoretical problems. The ^{57}Co decays include a line at 122 keV, which, with the ^{57}Co abundance 5 times solar, gives a flux that is marginally ruled out by existing observations⁵¹. The Gamma-Ray Observatory should provide a more definitive answer to this question. Observations of infrared Co and Fe lines also imply a $^{57}\text{Co}/^{56}\text{Co}$ ratio of 1–2 times the

solar value^{62,51}. Although some other source may be responsible for this luminosity, the uncertainties in the observations and models are probably still large enough for radioactivity to be a plausible source. The ⁵⁷Co source should eventually decay to the point where energy from ⁴⁴Ti decays dominates the ultraviolet through infrared luminosity at a level of $\sim 10^{36}$ erg s⁻¹ (ref. 58, fig. 2). The half-life of ⁴⁴Ti, 47 years, is long enough for radioactivity to continue to dominate some other potential luminosity sources, such as optically thick black-hole accretion or thermal emission from a neutron star, for decades.

Interaction with surroundings

The supernova affects its surroundings most directly through its radiation. At the time of shock breakout from the stellar surface, the photosphere is expected to be hot enough to emit ionizing radiation³. Narrow ultraviolet emission lines of highly ionized atoms, such as N III, N IV, N V, C III, O III and He II, were observed from SN1987A beginning ~ 80 days after the explosion⁵. The steady rise in the line fluxes with time could be attributed to the fact that it takes time for light to travel to various parts of the nebula before it can be seen by a distant observer. An electron density of $(1-4) \times 10^4$ cm⁻³ was deduced from the relative line intensities, a distance from the gas to the supernova of 200 light days from the line evolution, and a mass of 0.03–0.05 M_⊙ from the line strengths^{5,63}. Narrow lines of [O III] at 4,959 and 5,007 Å and other optical lines have also been observed^{64,65}; the slow evolution of the [O III] lines suggests an electron density $< 10^4$ cm⁻³ for this component⁶³. Detailed models including the time-dependent cooling and recombination of dense gas in a spherical shell around the supernova were able to reproduce many of the emission-line properties and suggested that the initial ionizing burst had a photon temperature in the range $(3-6) \times 10^6$ K (ref. 63).

More recent observations have shown that the emitting gas is in a ring as opposed to a shell. The evidence is the lack of emission inside the ring in an image with the Hubble Space Telescope⁶⁶ (Fig. 3) and the fact that the velocity field is consistent with an inclined, expanding (or contracting) ring⁶⁷. These results are from observations in optical emission lines, but it is likely that the ultraviolet lines are from the same region. Light-travel-time effects across an inclined ring in fact fit some of the

observed properties of the ultraviolet line evolution better than models with a spherical shell^{68,69}. The ring is inferred to have a radius of 6.3×10^{17} cm or 0.2 pc (for a LMC distance of 50 kpc) and is inclined to the plane of the sky by 43°. Using a simple model for the ultraviolet line emission, Panagia *et al.*⁶⁶ have deduced a distance of 51.2 ± 3.1 kpc to the supernova by combining the line evolution with the measured angular size. When realistic physical models are used for the line emission properties, this method holds promise for accurate distance determination. From velocity measurements, the ring is inferred to be expanding at 10.3 km s⁻¹ so that its radius/velocity age is 2×10^4 years (ref. 67), which is presumably the rough age of the red to blue supergiant transition.

Imaging of the supernova with the ESO New Technology Telescope shows larger-scale emission that suggests that the ring is part of a bipolar nebula⁷⁰. The nebula has been modelled on the assumption that it formed by the interaction of the fast wind from the blue supergiant progenitor star with the dense wind from a previous red supergiant phase^{71,72}. Luo and McCray⁷¹ assume that the shocked blue supergiant wind exerts a uniform pressure on a dense wind that has a factor 5 higher density in the equator than at the poles, and find the resulting shape to be comparable to that observed. The radial extent of the nebula in the polar direction is inferred to be 0.8 pc. The fact that the ring is so much more intense than the rest of the nebula suggests that a 'cusp' has formed in the wind interactions⁷², a property which has been found in models for planetary nebulae⁷³. The reason for the asymmetry in the red supergiant wind is unclear; it may be related to binary evolution, as has been suggested for similar structure observed in galactic bipolar nebulae⁷⁴. The projected major axis of the ring is at a position angle of $80^\circ \pm 5^\circ$ and, if the binary plane is the plane of rapid stellar rotation, the inner parts of the supernova envelope should be elongated along that position angle¹⁴. The observed position angle of the supernova major axis⁷⁵ is $20^\circ \pm 5^\circ$, which agrees with the position angle of supernova polarization measurements⁷⁶, so there is a discrepancy.

Outside the filamentary nebula, more diffuse emission has been observed that can be identified with dust-scattered light from the undisturbed red supergiant wind. Crotts and Kunkel⁶ estimate a mass-loss rate from the red supergiant of 3×10^{-6} M_⊙ yr⁻¹ over a timescale of $\sim 10^6$ years. The wind seems to end in a shell^{6,77,78} (Fig. 4), which can be identified with the position where the red supergiant wind is decelerated by the pressure of the surrounding medium⁷⁹; the shell radius is ~ 4 pc. Between this shell and the dense bipolar nebula, a structure known as Napoleon's hat⁷⁰ is present. Podsiadlowski *et al.*⁸⁰ suggest that this structure is the result of colliding winds from a binary progenitor system. In this model, the rings at a distance of 0.8 pc from the progenitor star are identified with the base of the dense colliding wind structure, as opposed to a bipolar nebula. Further imaging should allow us to distinguish between the models.

Beyond the 4-pc shell, there is a large region in front of the supernova that seems to be free of dust. This region may have been swept clean by the action of the main-sequence winds from the supernova progenitor star and from neighbouring massive stars. Dust clouds or shells are observed at 120 pc and 320 pc along the line of sight in front of the supernova, as revealed by circular light echoes from the dust^{81,82}. By following the structure of these echoes with time, it is possible to map out the three-dimensional structure of the interstellar medium close to the line of sight. The dust column density required to produce the echoes is comparable to the column density inferred from the reddening of the supernova. It is predicted that there should be an ultraviolet echo at the outer edge of the optical echo, produced by energetic radiation emitted close to the time of shock breakout⁸³. The brightness of the ultraviolet echo depends both on the unobserved initial flash from the supernova and on the poorly known ultraviolet scattering properties of dust grains.

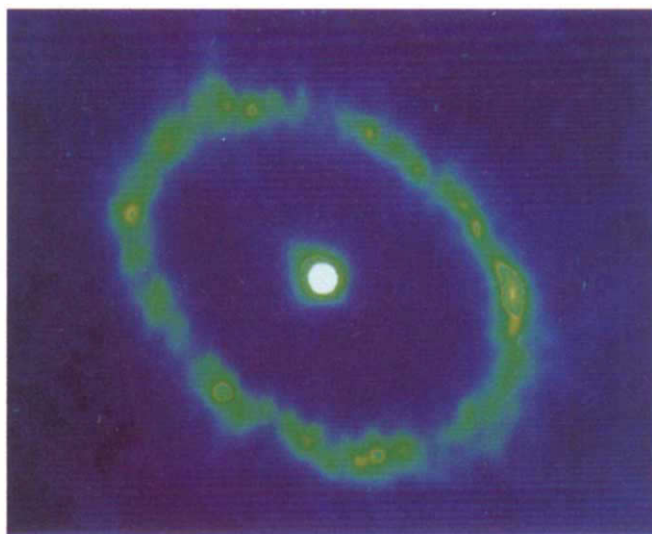


FIG. 3 An image of SN1987A (central star) and its surrounding ring with the Hubble Space Telescope Faint Object Camera in the light of the [O III] line at 5,007 Å on 24 August 1990. North is to the upper right corner and east is to the upper left corner. The major axis of the ring is 1.66 arcsec across, or 0.4 pc at a distance of 50 kpc (from ref. 66, courtesy of N. Panagia).

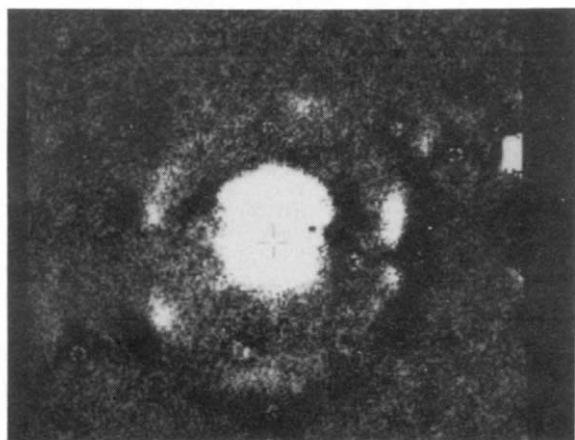


FIG. 4 Difference in optical continuum flux (March 1989 minus December 1989) showing the evolution of the stellar wind dust echo. The oval ring is ~ 20 arcsec across (from ref. 6, courtesy of A. Crotts).

Panagia and Gilmozzi⁸⁴ found evidence for such an echo in observations with the International Ultraviolet Explorer and deduced a cloud thickness of ~ 40 pc.

The radiative effects in the surrounding medium have given us an unusually complete picture for the environment of a supernova. The shock interaction with that environment is just beginning. The supernova explosion generates rapidly moving, cool gas that is in free expansion. It initially interacts with the stellar wind from the progenitor star, creating a high-pressure shell which gradually decelerates. The high-pressure shell is a likely source for the radio emission observed from SN1987A within the first few weeks after the explosion^{85,86}. This emission faded because of the decreasing density in the stellar wind, but reappeared⁸⁷ in July 1990 in an event that required a large increase in the efficiency of radio emission production. The timing and rough size (0.1 pc radius) of the event are consistent with the interaction of the supernova shock wave with the termination shock of the blue supergiant wind, where it is shocked by interaction with the red supergiant wind⁸⁸. The increase in radio efficiency may be related to turbulence and shock waves generated in the interaction region. In this model, the radio emission is predicted to decline when the shock front enters the relatively smooth region beyond the termination shock.

Relation to other supernovae and prospects

The main characteristic of the SN1987A light curve, a slow rise to maximum over a period of 80 days, has not been observed in other supernovae. This may be because SN1987A was fainter than typical type II supernovae at maximum by about an order of magnitude and because it is difficult to catch supernovae on their rise to maximum. SN1909A may have had a similar light curve, but there is only an upper limit before maximum light⁸⁹; its location in the outer part of a spiral galaxy, where the heavy element abundances are lower than solar, is consistent with the hypothesis that massive stars in metal-poor regions can explode as blue supergiants. Langer⁹⁰ has proposed that up to 40% of supernovae produced by the core collapse of single stars in regions with metallicities like that of the LMC may have blue supergiant progenitors.

At a distance of typical supernovae, SN1987A would currently be unobservable because of its low luminosity. Some type II supernovae, such as SN1980K (ref. 91), have recently been optically detected at an age of 5–10 years. These supernovae are also strong radio sources⁹², which implies circumstellar interaction with a dense red supergiant wind. The supernova gas is illuminated by energetic radiation from the interaction region and emits broad optical emission lines. These supernovae

indicate what might be expected once the supernova shock front in SN1987A reaches the old red supergiant wind. The supernova shock is predicted to reach the dense ring of gas between 2000 and 2005 (refs 71, 93). The shock wave driven into the dense gas should be radiative, giving rise to ultraviolet and optical emission lines. The interaction is also expected to produce strong X-ray emission and infrared emission from dust heated by the hot gas. The radio flux should increase, but may not be so strong because of the relatively small covering factor of the ring. The radio emission will increase as the shock front gradually interacts with the whole bipolar nebula.

A nearby supernova remnant where we can observe this type of interaction is Cassiopeia A, which has an age of ~ 300 years and is thought to be the explosion of a blue Wolf-Rayet star⁹⁴. Cas A shows both fast-moving knots, primarily made up of heavy elements, and quasi-stationary flocculi, made up of gas with a relatively normal composition except for an overabundance of He and N. It is plausible that the slow-moving gas is part of a shocked shell at a radius of 1.5 pc that was created by presupernova mass loss⁹⁵. The supernova shock is currently driving cooling shock waves into the denser shell gas. The fast-moving knots seem to be moving with constant velocity, and they cover a large velocity range; it is likely that they formed as a result of hydrodynamic instabilities during the supernova explosion, and it would be interesting to apply the instability models to this object, where the spatial structures can be observed. The abundance structure in the remnant gives evidence for macroscopic mixing of inhomogeneous layers⁹⁵.

Two supernova remnants in the LMC with ages of $\sim 10^3$ years also provide a constructive comparison. The remnant N132D shows an expanding ring of O with a velocity of $2,250 \text{ km s}^{-1}$ and a radius of 3 pc (ref. 96). It may be that the emission is from interaction with a circumstellar ring, and SN1987A will eventually show a similar phenomenon on a smaller scale. The remnant 0540–69 shows a central 50-ms pulsar with a surrounding synchrotron nebula and an optically emitting shell which was probably swept up by the pulsar nebula⁹⁷. The current pulsar power output is $1.5 \times 10^{38} \text{ erg s}^{-1}$ and the X-ray luminosity of the pulsar nebula is $3 \times 10^{36} \text{ erg s}^{-1}$ (ref. 25). The possibility that a comparable pulsar will emerge in SN1987A cannot yet be ruled out.

The nature of the central compact object remains the greatest mystery of SN1987A. Other puzzles include the large-scale asymmetries observed in the heavy element ejecta (Fe-group line emission), the supernova envelope (optical polarization), and the circumstellar medium ([O III] ring), which are in addition to the complex structures resulting from hydrodynamic instabilities. The continued expansion of the supernova will give a clearer view to the centre and will allow imaging of the supernova structure. SN1987A is currently making the transition between its supernova and supernova remnant phases. Although our detailed knowledge of the circumstellar medium lets us predict the coming interaction, more surprises are likely and will be most welcome. $\square$

Roger A. Chevalier is at the Department of Astronomy, University of Virginia, PO Box 3818, Charlottesville, Virginia 22903, USA.

1. Suntzeff, N. B., Phillips, M. M., Elias, J. H., De Poy, D. L. & Walker, A. R. *Astrophys. J.* **384**, L33–36 (1992).
2. Walborn, N. R. *et al. Astr. Astrophys.* **219**, 229–236 (1989).
3. Woosley, S. E. *Astrophys. J.* **330**, 218–253 (1988).
4. Shigeyama, T. & Nomoto, K. *Astrophys. J.* **360**, 242–256 (1990).
5. Fransson, C. *et al. Astrophys. J.* **336**, 429–441 (1989).
6. Crotts, A. P. S. & Kunkel, W. E. *Astrophys. J.* **366**, L73–77 (1991).
7. Arnett, W. D. *Astrophys. J.* **383**, 295–307 (1991).
8. Langer, N. *Astr. Astrophys.* **243**, 155–159 (1991).
9. Allen, D. A., Meikle, W. P. S. & Spyromilio, J. *Nature* **342**, 403–405 (1989).
10. Eastman, R. G. & Kirshner, R. P. *Astrophys. J.* **347**, 771–793 (1989).
11. Saio, H., Kato, M. & Nomoto, K. *Astrophys. J.* **331**, 388–393 (1988).
12. Hillebrandt, W. & Meyer, F. *Astr. Astrophys.* **219**, L3–6 (1989).
13. Podsiadlowski, Ph., Joss, P. C. & Rappaport, S. *Astr. Astrophys.* **227**, L9–12 (1990).
14. Chevalier, R. A. & Soker, N. *Astrophys. J.* **341**, 867–882 (1989).
15. Bionta, R. M. *et al. Phys. Rev. Lett.* **58**, 1494–1496 (1987).

16. Hirata, K. *et al.* *Phys. Rev. Lett.* **58**, 1490-1493 (1987).
17. Arnett, W. D., Bahcall, J. N., Kirshner, R. P. & Woosley, S. E. *Ann. Rev. Astr. Astrophys.* **27**, 629-700 (1989).
18. Burrows, A. *Astrophys. J.* **334**, 891-908 (1988).
19. Colgate, S. A. *Astrophys. J.* **163**, 221-230 (1971).
20. Chevalier, R. A. *Astrophys. J.* **346**, 847-859 (1989).
21. Michel, F. C. *Nature* **333**, 644-645 (1988).
22. Lin, D. N. C., Woosley, S. E. & Bodenheimer, P. H. *Nature* **353**, 827-829 (1991).
23. Blandford, R. D., Applegate, J. H. & Herquist, L. *Mon. Not. R. astr. Soc.* **204**, 1025-1048 (1983).
24. Sato, H. & Yamada, Y. *Prog. theor. Phys.* **85**, 541-546 (1991).
25. Seward, F. D. & Wang, Z.-R. *Astrophys. J.* **332**, 199-205 (1988).
26. Nomoto, K. & Tsuruta, S. *Astrophys. J.* **312**, 711-726 (1987).
27. Weaver, T. A. & Woosley, S. E. in *Supernova Spectra* (ed. Meyerott, R.) 15-32 (American Institute of Physics, New York, 1980).
28. Barbon, R., Capellaro, E. & Turatto, M. *Astr. Astrophys.* **135**, 27-31 (1984).
29. McCray, R. in *Supernovae and Stellar Evolution* (ed. Ray, A. & Velusamy, T.) 1-33 (World Scientific, Singapore, 1991).
30. Sunyaev, R. *et al.* *Nature* **330**, 227-229 (1987).
31. Dotani, T. *et al.* *Nature* **330**, 230-231 (1987).
32. Matz, S. M. *et al.* *Nature* **331**, 416-418 (1988).
33. Spyromilio, J. & Pinto, P. A. in *SN1987A and Other Supernovae* (ed. Danziger, I. J. & Kjær, K.) 423-429 (European Southern Observatory, Garching, 1991).
34. Fryxell, B., Müller, E. & Arnett, W. *Astrophys. J.* **367**, 619-634 (1991).
35. Hachisu, I., Matsuda, T., Nomoto, K. & Shigeyama, T. *Astrophys. J.* **358**, L57-62 (1990).
36. Müller, E., Fryxell, B. & Arnett, W. *Astr. Astrophys.* **251**, 505-514 (1991).
37. Herant, M. & Benz, W. *Astrophys. J.* (in the press).
38. Tueller, J. *et al.* *Astrophys. J.* **351**, L41-44 (1990).
39. Teegarden, B. J. *Adv. Space Res.* (in the press).
40. Colgate, S. A. in *Supernovae* (ed. Woosley, S. E.) 352-357 (Springer, New York, 1991).
41. Herant, M., Benz, W. & Colgate, S. A. *Astrophys. J.* (in the press).
42. Bussard, R., Burrows, A. & The, L.-S. *Astrophys. J.* **341**, 401-413 (1989).
43. Haas, M. R. *et al.* *Astrophys. J.* **360**, 257-266 (1990).
44. Cropper, M. *et al.* *Mon. Not. R. astr. Soc.* **231**, 695-722 (1988).
45. Jeffery, D. J. *Astrophys. J.* **375**, 264-287 (1991).
46. Yamada, S. & Sato, K. *Astrophys. J.* **358**, L9-12 (1990).
47. Kozma, C. & Fransson, C. *Astrophys. J.* (in the press).
48. Fransson, C. & Chevalier, R. A. *Astrophys. J.* **322**, L15-20 (1987).
49. Spyromilio, J. *et al.* *Mon. Not. R. astr. Soc.* **248**, 465-478 (1991).
50. Menzies, J. W. in *SN1987A and Other Supernovae* (ed. Danziger, I. J. & Kjær, K.) 209-215 (European Southern Observatory, Garching, 1991).
51. Danziger, I. J., Bouchet, P., Gouffes, C. & Lucy, L. in *SN1987A and Other Supernovae* (ed. Danziger, I. J. & Kjær, K.) 217-224 (European Southern Observatory, Garching, 1991).
52. Spyromilio, J. & Graham, J. R. *Mon. Not. R. astr. Soc.* (in the press).
53. Roche, P. F., Aitken, D. K., Smith, C. H. & Jones, S. D. *Nature* **337**, 533-535 (1991).
54. Lucy, L. B., Danziger, I. J., Gouffes, C. & Bouchet, P. in *Supernovae* (ed. Woosley, S. E.) 82-94 (Springer, New York, 1991).
55. Kozasa, T., Hasegawa, H. & Nomoto, K. *Astr. Astrophys.* **249**, 474-482 (1991).
56. Bouchet, P., Danziger, I. J. & Lucy, L. B. *Astr. J.* **102**, 1135-1146 (1991).
57. Suntzeff, N. B., Phillips, M. M., De Poy, D. L., Elias, J. H. & Walker, A. R. *Astr. J.* **102**, 1118-1134 (1991).
58. Woosley, S. E., Pinto, P. A. & Hartmann, D. *Astrophys. J.* **346**, 395-404 (1989).
59. Woosley, S. E. & Hoffman, R. D. *Astrophys. J.* **368**, L31-34 (1991).
60. Kumagai, S., Shigeyama, T., Hashimoto, M. & Nomoto, K. *Astr. Astrophys.* **243**, L13-16 (1991).
61. Sunyaev, R. A. *et al.* *Soviet Astr. Lett.* **16**, 171-176 (1990).
62. Varani, G. F., Meikle, W. P. S., Spyromilio, J. & Allen, D. A. *Mon. Not. R. astr. Soc.* **245**, 570-576 (1990).
63. Lundqvist, O. & Fransson, C. *Astrophys. J.* **380**, 575-592 (1991).
64. Wampler, E. J. & Richichi, A. *Astr. Astrophys.* **217**, 31-34 (1989).
65. Wang, L. *Astr. Astrophys.* **246**, L69-72 (1991).
66. Panagia, N., Gilmozzi, R., Macchetto, F., Adorf, H.-M. & Kirshner, R. P. *Astrophys. J.* **380**, L23-26 (1991).
67. Crotts, A. P. S. & Heathcote, S. R. *Nature* **350**, 683-685 (1991).
68. Lundqvist, P. in *SN1987A and Other Supernovae* (ed. Danziger, I. J. & Kjær, K.) 607-615 (European Southern Observatory, Garching, 1991).
69. Felten, J. E. & Dwek, E. in *SN1987A and Other Supernovae* (ed. Danziger, I. J. & Kjær, K.) 569-574 (European Southern Observatory, Garching, 1991).
70. Wampler, E. J. *et al.* *Astrophys. J.* **362**, L13-16 (1990).
71. Luo, D. & McCray, R. *Astrophys. J.* **379**, 659-662 (1991).
72. Wang, L. & Massali, P. A. *Nature* **355**, 58-61 (1992).
73. Kahn, F. D. & West, K. A. *Mon. Not. R. astr. Soc.* **212**, 837-850 (1985).
74. Morris, M. *Astrophys. J.* **249**, 572-585 (1981).
75. Papaliolios, C. *et al.* *Nature* **338**, 565-566 (1989).
76. Bailey, J. *Proc. astr. Soc. Austr.* **7**, 405-411 (1988).
77. Bond, H. E., Gilmozzi, R., Meakes, M. G. & Panagia, N. *Astrophys. J.* **354**, L49-52 (1990).
78. Sparks, W. B., Paresce, F. & Macchetto, D. *Astrophys. J.* **347**, L65-68 (1990).
79. Chevalier, R. A. & Emmering, R. T. *Astrophys. J.* **342**, L75-78 (1989).
80. Podsiadlowski, Ph., Fabian, A. C. & Stevens, I. R. *Nature* **354**, 43-46 (1991).
81. Suntzeff, N. B. *et al.* *Nature* **334**, 135-138 (1988).
82. Crotts, A. P. S. *Astrophys. J.* **333**, L51-55 (1988).
83. Chevalier, R. A. & Emmering, R. T. *Astrophys. J.* **331**, L105-108 (1988).
84. Panagia, N. & Gilmozzi, R. in *SN1987A and Other Supernovae* (ed. Danziger, I. J. & Kjær, K.) 575-594 (European Southern Observatory, Garching, 1991).
85. Turtle, A. J. *et al.* *Nature* **327**, 38-40 (1987).
86. Chevalier, R. A. & Fransson, C. *Nature* **328**, 44-45 (1987).
87. Staveley-Smith, L. *et al.* *Nature* **355**, 147-149 (1992).
88. Chevalier, R. A. *Nature* (in the press).
89. Young, T. R. & Branch, D. *Nature* **333**, 305-306 (1988).
90. Langer, N. *Astr. Astrophys.* **252**, 669-688 (1991).
91. Fesen, R. A. & Becker, R. H. *Astrophys. J.* **351**, 437-442 (1990).
92. Weiler, K. W., Sramek, R. A., Panagia, N., van der Hulst, J. M. & Salvati, M. *Astrophys. J.* **301**, 790-812 (1986).
93. Chevalier, R. A. & Liang, E. P. *Astrophys. J.* **344**, 332-340 (1989).
94. Fesen, R. A., Becker, R. H. & Blair, W. P. *Astrophys. J.* **313**, 378-388 (1987).
95. Chevalier, R. A. & Kirshner, R. P. *Astrophys. J.* **233**, 154-162 (1979).
96. Lasker, B. M. *Astrophys. J.* **237**, 765-768 (1980).
97. Kirshner, R. P., Morse, J. A., Winkler, P. F. & Blair, W. P. *Astrophys. J.* **342**, 260-271 (1989).
98. Axelrod, T. S. *Lawrence Livermore National Lab. preprint UCRi 52994* (1980).

ACKNOWLEDGEMENTS. I am grateful to C. Fransson, P. Lundqvist, D. Luo and S. Woosley for comments, to W. D. Arnett, N. Panagia and A. Crotts for figure material, and to the NSF for support. Data for Fig. 4 were obtained at Las Campanas Observatory, operated by the Observatories of the Carnegie Institution of Washington.

ARTICLES

Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix

Oxana Ibraghimov-Beskrovnaya, James M. Ervasti, Cynthia J. Leveille, Clive A. Slaughter*, Suzanne W. Sernett & Kevin P. Campbell†

Howard Hughes Medical Institute and Department of Physiology and Biophysics, University of Iowa College of Medicine, Iowa City, Iowa 52242, USA
* Howard Hughes Medical Institute, Biopolymer Facility, University of Texas Southwestern Medical Center, Dallas, Texas 75235, USA

The primary sequence of two components of the dystrophin-glycoprotein complex has been established by complementary DNA cloning. The transmembrane 43K and extracellular 156K dystrophin-associated glycoproteins (DAGs) are encoded by a single messenger RNA and the extracellular 156K DAG binds laminin. Thus, the 156K DAG is a new laminin-binding glycoprotein which may provide a linkage between the sarcolemma and extracellular matrix. These results support the hypothesis that

the dramatic reduction in the 156K DAG in Duchenne muscular dystrophy leads to a loss of a linkage between the sarcolemma and extracellular matrix and that this may render muscle fibres more susceptible to necrosis.

THE Duchenne muscular dystrophy (DMD) gene encodes dystrophin, a large membrane cytoskeletal protein found in skeletal muscle^{1,2}. Dystrophin is localized to the inner surface of the sarcolemma in normal skeletal muscle but is absent in skeletal muscle of DMD patients, mdx mice and xmd dogs³⁻⁶. Biochemical studies have demonstrated that dystrophin is tightly linked to a large oligomeric complex of sarcolemma glyco-

† To whom correspondence should be addressed.

proteins^{7,8}. A dystrophin-glycoprotein complex (DGC) has been isolated and shown to consist of cytoskeletal (relative molecular mass 59,000 (M_r 59K) and dystrophin), transmembrane (50K, 43K, 35K and 25K) and extracellular (156K) components⁹. A model for the organization of the DGC has also recently been proposed⁹. The membrane organization of the dystrophin-glycoprotein complex⁹ and the high density of dystrophin in the sarcolemma membrane^{10,11} suggest that this complex could have an important structural role in skeletal muscle, but its exact function remains unknown. Furthermore, the absence of dystrophin in skeletal muscle from mdx mice and DMD patients leads to a dramatic loss of all of the components of the dystrophin-glycoprotein complex^{8,12}. Because abnormal expression of the dystrophin-associated glycoproteins (DAGs) may have a crucial role in molecular pathogenesis in DMD, it is important to identify the normal function of the dystrophin-glycoprotein complex in skeletal muscle.

We report here the complete amino-acid sequence of 43K and 156K dystrophin-associated glycoproteins deduced from the cDNA sequence. Both components of the dystrophin-glycoprotein complex are encoded by the same 5.8-kilobase (kb) mRNA. Post-translational modification of a 97K precursor protein results in two mature proteins: 43K DAG and 156K DAG. Sequence analysis of cDNAs reveals an open reading frame encoding a precursor polypeptide with no significant sequence similarity with any known proteins. The N-terminal portion of the precursor polypeptide is processed into the mature 156K DAG with a putative protein core of about 56K with potential attachment sites for O-linked carbohydrates. The C-terminal portion of the precursor polypeptide is processed into the mature 43K DAG with potential N-glycosylation sites, a single transmembrane domain and a 120 amino-acid long cytoplasmic tail. Northern and western blot analyses have demonstrated that the 43K DAG and 156K DAG are expressed in both muscle and non-muscle tissues. The specific mRNA for the 43–156K DAG is expressed at normal amounts in mdx and DMD skeletal muscle, whereas both glycoproteins are greatly reduced in dystrophin-deficient muscle. The dramatic decrease in the 43K DAG occurs in the absence of dystrophin in muscle but not in other tissues of mdx mice. Finally, we show that the 156K DAG binds laminin, a well characterized component of the extracellular matrix. Thus, our results demonstrate that the 156K DAG is a novel laminin-binding glycoprotein and suggest that the function of the dystrophin-glycoprotein complex is to link the subsarcolemmal cytoskeleton to the extracellular matrix. We propose to name the 43–156K dystrophin-associated glycoprotein 'dystroglycan' because of its identification through dystrophin and its extensive glycosylation.

Cloning and sequence analysis

A λ gt11 cDNA expression library from rabbit skeletal muscle was screened with affinity-purified guinea pig polyclonal antibodies specific for the 43K DAG. One cDNA clone with a 0.6-kb insert designated R43-A was found and its sequence revealed one open reading frame (Fig. 1a,b). Overlapping clones covering the entire coding region of the mRNA were isolated by rescreening of rabbit skeletal muscle cDNA libraries using radiolabelled cDNA probes (Fig. 1a).

The 4,200-nucleotide cDNA sequence contains a 2,685 nucleotide open reading frame coding for a polypeptide of 895 amino acids with a calculated M_r of 97,029 (Fig. 1b). The first 29 amino acids of the open reading frame are predominantly hydrophobic and probably represent a signal peptide (Fig. 1b,c). Hydropathy analysis identified a single continuous region of 24 amino acids close to the C terminal showing characteristics of a transmembrane domain (Fig. 1c). Four potential N-linked glycosylation sites and numerous potential phosphorylation sites are found in the 97K polypeptide. No significant sequence homology was detected in the NBRF database between any

known proteins and the predicted amino-acid sequence of the 97K polypeptide.

Biochemical characterization of the dystrophin-glycoprotein complex has demonstrated that the 43K DAG has hydrophobic properties characteristic of a transmembrane protein and contains Asn-linked oligosaccharides⁹. These properties are consistent with the predicted sequence of the C-terminal half of the 97K polypeptide which contains a potential transmembrane domain and three out of four potential sites for N-glycosylation. The C-terminal origin of 43K DAG was confirmed using an antibody raised in a rabbit against a synthetic peptide corresponding to the 15 C-terminal amino-acid residues of the deduced sequence. This anti-peptide antibody specifically recognized the 43K DAG (Fig. 2b). In addition, peptide sequence determined directly from the 43K DAG matched residues 783–793 of the deduced amino-acid sequence of the 97K polypeptide (Fig. 1b).

N-terminal domain of the 97K precursor

To identify the N-terminal domain of the 97K precursor polypeptide, antibodies to different regions of the 97K precursor polypeptide were produced by expressing several overlapping cDNAs encoding different regions in the 97K precursor polypeptide. The pGEX vectors for the expression of foreign sequences as glutathione S-transferase (GST) fusion proteins in *Escherichia coli* cells were used to examine specific regions in the 97K precursor protein, corresponding to the 43K DAG, the N-terminal half of the 97K precursor and to sequence overlapping both the N- and C-terminal halves (Fig. 2a).

Monospecific antibodies to each fusion protein were affinity purified from the sheep antiserum raised against purified DGC using immobilized fusion proteins. Affinity-purified antibodies were then tested using each fusion protein and purified DGC (Fig. 2b). Consistent with the C-terminal domain encoding the 43K DAG, antibodies to FP-A (fusion protein-A) and FP-C specifically stained both bands of 43K DAG doublet (Fig. 2b). But antibodies to FP-B stained the 43K DAG and the 156K DAG components of DGC (Fig. 2b). Thus, a second product of 97K precursor polypeptide seems to be the 156K DAG. In accordance with this supposition, antibodies to FP-D stain only 156K DAG (Fig. 2b). Therefore, post-translational processing of 97K precursor polypeptide gives rise to two components of DGC: 43K DAG and 156K DAG. Biochemical studies have demonstrated that 156K DAG is not an integral membrane protein and contains N-linked and O-linked glycosylation⁹. These properties are consistent with the predicted N-terminal half of the 97K precursor which does not possess any hydrophobic region, has one potential N-linked glycosylation site and many potential O-glycosylation sites. Although generation of the 43K and 156K DAG by proteolysis during the preparation of muscle membranes cannot be ruled out, we have never seen a larger protein react with antibodies to the 43K DAG or 156K DAG and the 156K DAG is seen in direct SDS extracts of muscle⁸. Further analysis of 43–156K-specific mRNA translation is needed to show that the cleavage of the 97K precursor occurs during its synthesis and is not occurring during the isolation procedure. The precise point of cleavage in 97K precursor protein used in the formation of the 43K DAG has not been determined because the N terminal of 43K DAG was blocked and thus could not be sequenced. Among the putative cleavage sites, Arg 457 is the best fitting according to consensus sequence for post-translational processing of protein precursors at monobasic sites¹³.

Expression of 43–156K DAG

A prominent 5.8-kb transcript was detected in mRNA from rabbit skeletal muscle, cardiac muscle and lung (Fig. 3a). A weaker hybridizing transcript of the same size was found in brain (Fig. 3a). Northern blot analysis with total RNA from liver, kidney, diaphragm and stomach also detected a 5.8-kb mRNA in all these tissues (data not shown). Thus, the 5.8-kb

transcript for the 43–156K DAG is present in various muscle and non-muscle tissues, most probably originating from the same gene. The 43–156K DAG in muscle and non-muscle tissues was identified using immunoblots of membranes from different tissues and affinity-purified antibodies to FP-B (43–156K specific) (Fig. 3*b*). The 43K DAG was detected in isolated membranes from skeletal muscle, brain, cardiac muscle and lung (Fig. 3*b*). The 156K DAG was detected in skeletal and cardiac muscle membranes, but has a slightly lower M_r in cardiac membranes (Fig. 3*b*). In brain and lung membranes the M_r of the '156K' DAG reactive protein was ~120K. The variability in M_r for the '156K' reactive protein may be due to differential glycosylation of the core protein in muscle versus non-muscle tissues.

Analysis of dystrophic muscle

RNA blot analysis of skeletal muscle mRNA from control and mdx mice of different ages using cDNA probe R43-B to 43–156K mRNA revealed no reduction of 43–156K DAG mRNA in mdx mice versus control mice (Fig. 4*a*). But as previously observed^{8,12}, the 43K DAG is greatly reduced in mdx skeletal muscle membranes (Fig. 4*b*). Thus, the absence of dystrophin causes no change in the mRNA for the 43–156K DAG but leads to dramatic reductions in the amount of the 43K DAG and 156K DAG in skeletal muscle. Analysis of mRNA from control and DMD skeletal muscle also showed no difference in 43–156K DAG mRNA expression (Fig. 5*a*). In agreement with findings in mdx mouse muscle, indirect immunofluorescence analysis of cryosections from normal and DMD skeletal muscle with 156K-specific (anti FP-D) and 43K-specific (anti FP-A) antibodies demonstrated a drastically reduced density of 43K DAG and 156K DAG in skeletal muscle of a DMD patient (Fig. 5*b*). Thus the 43–156K DAG-encoding gene is transcribed and specific mRNA is still present at the normal amount in dystrophic muscle, but the 43K DAG and 156K DAG are greatly reduced in dystrophic muscle.

Because the 43–156K DAG is expressed in non-muscle tissues we also examined expression of 43K DAG in the non-muscle tissues of control and mdx mice. The 156K DAG could not be tested because polyclonal antibodies to the protein core of rabbit 156K DAG described above do not crossreact with the 156K DAG in mouse muscle. Immunoblot analysis of brain and kidney membranes from control and mdx mice, stained with polyclonal anti FP-A antibodies (43K specific), revealed no reduction in the amount of 43K DAG in these mdx tissues (Fig. 4*b*). Thus, the dramatic reduction of the 43K DAG that is found in mdx mice seems to be restricted to skeletal muscle and is not found in non-muscle tissues.

156K DAG binds laminin

The high density of dystrophin in isolated sarcolemma and the membrane organization of the DGC have suggested a structural function for the DGC^{9,11}. That the 156K DAG is an extracellular component of the dystrophin-glycoprotein complex suggests it may interact with the extracellular matrix. To test for the association of the 156K DAG with the extracellular matrix, rabbit skeletal muscle surface membranes and pure DGC were electrophoretically separated, transferred to nitrocellulose membranes and overlaid with ¹²⁵I-labelled laminin. A single laminin-binding band, corresponding to the 156K DAG, was detected in surface membranes and purified DGC (Fig. 6*a*). Binding of ¹²⁵I-labelled laminin to the 156K DAG was significantly decreased by a 1,000-fold excess of unlabelled laminin demonstrating the specificity of ¹²⁵I-labelled laminin binding to the 156K DAG (Fig. 6*a*). ¹²⁵I-labelled fibronectin did not label the 156K DAG or any other component of the DGC, nor did a 1,000-fold excess of non-radioactive fibronectin have any effect on the binding of ¹²⁵I-labelled laminin to the 156K DAG (data not shown). The interaction of 156K DAG with laminin was also shown by coimmunoprecipitation of laminin and 156K DAG (Fig. 6*b*).

Anti-laminin antibodies did not precipitate the 156K DAG from alkaline extracts of rabbit skeletal muscle surface membranes (Fig. 6*b*, lane 3). This result was consistent with our observation (not shown) that the surface membranes used were devoid of laminin, merosin¹⁴ or S-laminin¹⁵ as detected on immunoblots using specific antibodies¹⁶. But anti-laminin antibodies effectively precipitated the 156K DAG from alkaline extracts that had been preincubated with exogenously added laminin (Fig. 6*b*, lane 7). These results suggest that the 156K DAG specifically binds laminin and may mediate interaction of DGC with extracellular matrix.

Discussion

We report here the nucleotide and deduced amino-acid sequences for two components of the dystrophin-glycoprotein complex. We show that the 43K DAG and 156K DAG are encoded by the same mRNA and that post-translational processing of the 97K precursor polypeptide results in two mature proteins: 43K DAG and 156K DAG. The mature 43K DAG contains a single transmembrane domain, three potential sites for N-glycosylation and a 120 amino-acid long cytoplasmic tail whereas the 156K DAG contains no transmembrane sequences,

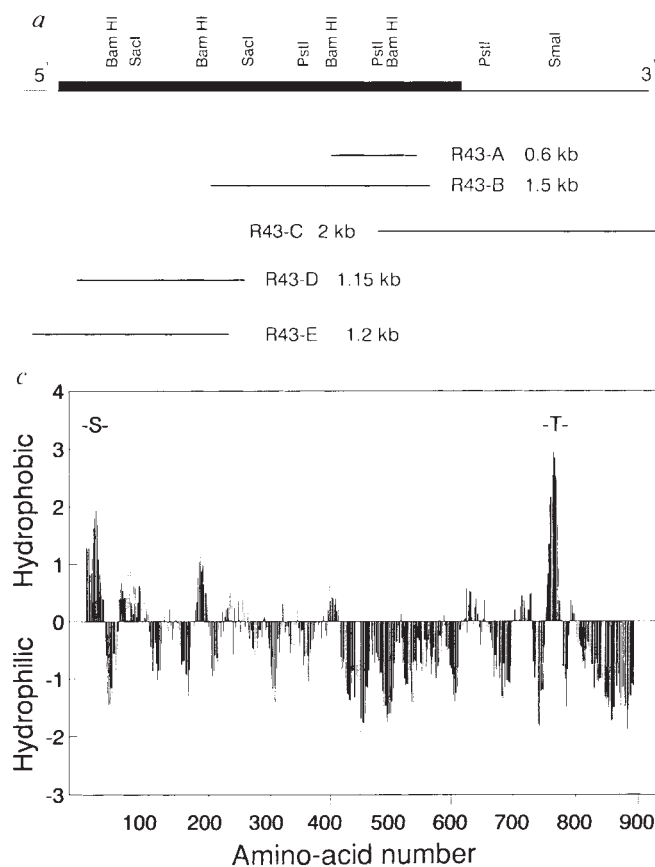


FIG. 1 Cloning and sequence analysis of cDNAs encoding a precursor for the 43–156K DAG. *a*, Restriction map and overlapping cDNA clones encoding the precursor protein for the 43–156K DAG. The protein-coding region is indicated by a solid box. *b*, The nucleotide and predicted amino-acid sequence (single-letter code) of precursor protein for the 43–156K DAG. The putative signal peptide is underlined and the predicted transmembrane domain is double underlined. A sequence matching that of a peptide isolated from the mature 43K DAG is indicated by a dashed line. *, Potential asparagine-linked glycosylation sites. *c*, Hydropathy plot of the 43–156K DAG precursor protein. Hydropathy profile of the 43–156K DAG precursor protein computed according to ref. 29; the window size is 19 residues plotted at one-residue intervals. T, Position for the predicted transmembrane domain; S, proposed signal sequence.

METHODS. Affinity-purified guinea pig polyclonal antibodies to the 43K DAG were prepared as previously described⁹ and used to screen 2×10^6 clones

b -169 GGCTGCTTTTCAGGAAGATAAAGCTTTTAAGGCTGCCTAACACTAGAAG -121
 GAGAGGCTCTCGATGCTCTGGGATGGAGCAGGTGTGCAGAGGGTGAAGGACCCGGCTCTGGGATCAAGTCACTTGCTTGCTTCTAGCAAGATCTCGCTTGAGCGAACTTGGCCCTGGG -1

ATGAGGATGTCTGTGGGCTTTCTACTGCTGCTCCCTTGTGGGGAGGACATTCTCTCTCTCTGTGTGGCGTGGCTCAGTCCCATTTGGCCAGCGAACCTCGGAGGCTGTGAG 120
M R M S V G L S L L L P L W G R T F L L L L C V A V A Q S H W P S E P S E A V R 40

GACTGGGAGAACCAGCTGGAGGCTCCATGCACCTGTGCTCTCAGACCTGCACGAAGCCCTTCCACAGTGGTTGGCATTCTGATGGCAGCGCTGTTTGGGCGCTCGTTTCGAGTG 240
 D W E N Q L E A S M H S V L S D L H E A L P T V V G I P D G T A V V G R S F R V 80

ACCATTCCACAGATTAAATTTGGCTCCAGTGGAGAAGTCATCAAGGTATCCACGGCAGGAAGAGGTTTGGCCATCGTGGCTGCATTGGGATCCACAGAGCCACACCTGGAGGGCCCTT 360
 T I P T D L I G S S G E V I K V S T A G K E V L P S W L H W D P Q S H T L E G L 120

CCGCTGGACACGACAAGGTTGTGATTACATCTCAGTGAGCGCTGCACAGCTGGATGCCAACGGAAGCCACATCCCTCAGACCTCCAGTGTGTTTCCATCGAGGCTACCCCGGAAGAC 480
 P L D T D K G V H Y I S V S A A Q L D A N* G S H I P Q T S S V F S I E V Y P E D 160

CACAGTGAGCCGACGTGTGTGGGCGGCTCTCCAGACCTGGGCGAGGCGGCGCTGTGCTGTGCTGCCAGGAGCGGTGACCGTCTTGACCGTGATTCTGGATGCCGATCTCACC 600
 H S E P Q S V R A A S P D L G E A A A S A C A A E E P V T V L T V I L D A D L T 200

AAGATGACTCCGAAGCAGAGGATCGACCTCTGCACAGGATGCAGAGCTTCTCGGAGTGGAGCTCCACAACATGAAGTTGGTCCGGTGGTGAATAACAGACTGTTTGATATGTCTGCC 720
 K M T P K Q R I D L H R M Q S F S E V E L H N M K L V P V V N N R L F D M S A 240

TTTATGGCCGGCCCCGAAACGCCAAAGTGGTAGAAGCGGGCCCTGCTCTCTGGAAGCTGGGCTGCTCCCTGAACAGAACAGTGTGCTGACATTGCGCGCTGGAGGCCCTT 840
 F M A G P G N A K K V V E N G A L L S W K L G C S L N Q N S V P D I R G V E A P 280

GCCAGGAGGGCACTATGTCTGCCAGCTTGGTACCCTGTGGTGGTGGCAGTATGCCAACGAAGCACCTCTCCCAAGCGTATCCGAAGGAGATCCATGCCACACCCACCT 960
 A R E G T M S A Q L G Y P V V G W H I A N K K P P L P K R I R R Q I H A T P T P 320

GTCAGTGCATTGGGCCCCAACACGGCCATCCAGGAGCGCGCTCCAGGATCGTCCCTACCCCACTTCTCCAGCCATTGCTCTCCACAGAGAGCATGGCTCTCCAGTCCAGGAT 1080
 V T A I G P P T T A I Q E P P S R I V P T P T S P A I A P P T E T M A P P V R D 360

CCTGTTCTGGGAAGCCACGGTCCACTCGGACTCGAGGTGCCATTATTCAGACCCCAACCTAGGCCCATCCAGCCACTCGGGTGTGAGACGCTGGCACCCTAGTTTCTGGCCAG 1200
 P V P G K P T V T T R T R G A I I Q T P T L G P I Q P T R V S D A G T V V S G Q 400

ATTCGTCAACGGTGACCATCTCGGTACGTGGAGCCACAGCAGTTGCCACCCCTCCCAACTACAACCAAAAGCCACGAGTGTCCACACCAAAACAGCAACGCCCTTCAACGGAC 1320
 I R A T V T I P G Y V E P T A V A T P P T T T T K K P R V S T P K P A T P S T D 440

TCCTCAGCCACACGACTCGCAGGCCAACCAAGAGCCACGGACACCCAGGCGGTGCCAGGGTCACTAAAGCTCCCATCACCAGGCTGGAGACGGCCCTCCCACTACTCGTATC 1440
 S S A T T T R R P T K K P R T P R P V P R V T T K A P I T R L E T A S P P T R I 480

CGCACCCACACGCGGGTGGCCCGGGGAGAACCAACACGCGCCAGAGCTCAAGAACACATCGACAGGGTGGACGCTGGGTGGGCACTTCTTGAAGTGATCCCATCT 1560
 T T T S G V P R E L K N H I D R V D A W V G T F A G G K I P S 520

GATACCTTACGACAAGGAGTACCACACCGACAAGCTCAAGCTGACCTGAAGCTGCGAGAGCAGCAGTGGTGGGCGAGAAGTCTGGTGCAGTTTCAACAGCAACAGCCAGCTC 1680
 D T F Y D K E D T T T D K L K L T L K L R E Q Q L V G E K S W V Q F N S N S Q L 560

ATGTATGGCTGCCGACAGCAGCCACGTGGGCAACACGAGTATTTTCATGCATGCCACAGACAAGGAGGCTGTCCGCGTGGATGCCCTTGAGATCCATGTCCACAAGCGCCCTCAA 1800
 M Y G L P D S S H V G K H E Y F M H A T D K G G L S A V D A F E I H V H K R P Q 600

GGGACAAGCTCCTGCTGTTTCAAAGCAAGTTCGTGGTGACCCAGCGCCAGTGGTGAATGACATCCACAAGAAGATTGCCCTGGTGAAGAAGCTGGCTTGGCTTGGGATCGC 1920
 G D K A P A R F K A K F V G D P A P V V N D I H K K I A L V K K L A F A F G D R 640

AATTGCAGCACCCTCACCCTGCAGAACATCACCCGCGCTCCATTGTGGTGGAGTGGACCAACAACACACTGCCCTGGAGCCCTGCCCAAGGAGCAGATCACGGGCTGAGCCGACG 2040
 N* C S T V T L Q N* I T R G S I V V E W T N* N T L P L E P C P K E Q I T G L S R R 680

ATCGCCGAGGACAACGGGACGCTCGGCCAGCCTTCAACATGCCCTGGAGCCTGACTTTAAGGCCACGAGCATGCCATAACGGGCTCTGSCAGTTGTGSCACTTGCAGTTTATCCCC 2160
 I A E D N G Q P R P A F T N A L E P D F K A T S I A I T G S G S C R H L Q F I P 720

GTGGCACCCTGGGATCCGCTCTCGTGACACCCACGAGGTGCCAGAGGCCCGGAGAGGAGTGGAGTACGCTACCTACACAGTCATTCCGGCTGTGGTGGTG 2280
 V A P P G I P S V I P T P T E V P D R D P E K S S E D D V Y L H T V I P A V V V 760

GCGGCCATCTGCTCATTTGCTGGCATTTGCCATGATCTGCTACCGCAAGAAGCGGAAGGCAAGCTCACCTGGAGGACCGCCACCTTCAAGAAGGGGCTGCCATCTTT 2400
A A I L L I A G I A I A M I C Y R K K R K G K L T L E D Q A T F I K G V P I I F 800

CGAGACGAGCTGGAGACTCCAAGCCCCCGCTCTCCAGCATGCCGCTGATCTCGAGAGGAGAAGGCTCCCTTCCCCCCCCAGAGTATCCAGCCAGAGCGTCCCGAGACACG 2520
 A D E L D D S K P P P S S S M P L I L Q E E K A P L P P P E Y P S Q S V P E T T 840

CCTCTGAACAGGACACTGTGGGGAGTACACGCCCTTGGGATGAGGATCCCAACGCGCTCCCTACCAGCCCCCCCCACCTTCCAGCCCCGATGGAGGCAAGGGCTCCCGTCCC 2640
 P L N Q D T V G E Y T P L R D E D P N A P P Y Q P P P P P T A P M E G K G S R P 880

AAGAATCATGCCCTTACCGGTACCCCTCTTATGTTCCCTTAAACCCACAAGCGCTGGTGGAGCAGGGTAGGGCAGGGGCTGGGACAACACAGTGTGTTGCTGGAGCCCG 2760
 K N M T P Y R S P P P Y V P P STOP 895

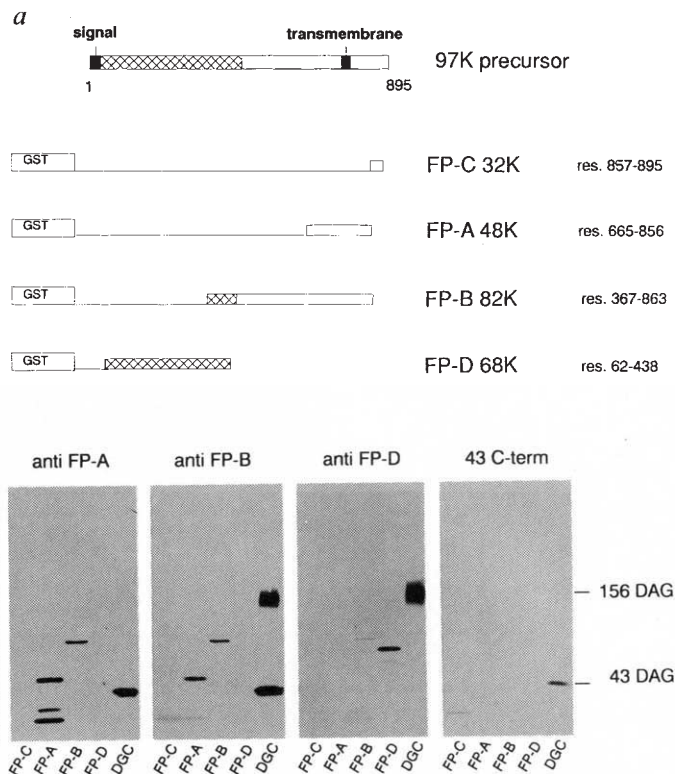
GTGGCCGAGACCATCGCCACTGGGCGTGACACAGACCTAGCACACTGGCACACGGGGCTGGACAAGCCCGCCCTCTCTGCTCTCCCAACCCCAAGCAGCTGGAGAGACT 2880
 TTGGGACTTTTTTATTTTTTTTGGCTAACAGCTTTTGTGTTTTCATAGAAAATCTTCGCTGCGTTTGTATGGCTGGCTCGGAAGCACCATTGGAGTAGAGGTAGAGGGAG 3000
 GGAGCGAGGAGCGGTGGTGAAGTGCAGGCACTGTGGGACGCCCCCGCTCTCTGCGTTTGGCTTAAACACTAAGTACTGTTTTTCTATTACGTTGTCTAGCTGACAGGATG 3120
 TAACATGGAAACAGTAAAGTAAATTCAGAGGCTTTCAGAGTAAAGTAAAGTTTACATTTTAACTGCTGTTTACCTAAACTGTATGATAATTTTGGTGGGTATGG 3240
 GGAATTGCTGCTAAATAAGCTCCCAAGGTTTCAAACCTTAAGAGAAGCAAGGACAGTATTTTATCAAAGGAATCTATTTTTCACACTATGTCAACTTGGTGTGCTGTA 3360
 TATCCAGAGCCGATCGGGGGCTCTGGCCCTGGCTCAGGGGCAAGGCTGGTGTGGTGTGGTCTGCTGCTGCTGTTGGCAGGAGTGGAAAGTGGAGGGGCTCTCGGGCGGTGA 3480
 CATCCCCACCTCCACCCATGATGCTAGTGGCCACCAAGGGGCTTCAATTCCATGGGAAAGGAGCTCCAGAGGCAAGTGGTGGCGGTGGCCCCACCCCGGGTGTCCAAGGTG 3600
 GGCCAGCTGCTCGTGGGGGCTTGGGAGGCTGACACACATCGACCTGTTTCTTTCACCTTTTATTTTTTTTCCCAACCCCTCAAAAGGATTATCACGTTTGTG 3720
 AAACACTCAGTGGGGGACATTTGTGAAGATGCAATATTTTATGTCATGTGATGCTCTTCTCCTCAGTGTGACCTTGGCCACTTGTGCCAACAGTCCACAGCCCCCGCCGATCCACCC 3840
 CATCCCTCTCTCTGGCGCTCCCGTCCAGGCTTGGGCTGAACGACTGGAAGAGGCTGTTGGCTGGGGAGGAGTCCACCAATAGTTCATAGTAAACAATCTGTGGGCTCTCAAG 3960
 CTAATTTTTTACTAAGTTTTTATACGCTCAAAATGTTTTATTAATAAATAGATTAATAATGGTGATGC 4031

of λ gt11 expression library. Clone R43-A with a length of 600 base pairs (bp) was isolated from a random primed adult rabbit skeletal muscle λ gt11 library²⁸ by immunoscreening. An oligo-dT-primed rabbit skeletal muscle cDNA library in λ zapII (Stratagene) was screened at high stringency with a ³²P-labelled cDNA insert from the R43-A clone (random primed labelling kit, Boehringer Mannheim). Clone R43-B overlaps R43-A and extends ~1 kb in the 5' direction. Further clones were isolated from λ gtII libraries³⁰: R43-D,

random-primed λ gt11 library; R43-C, oligo-dT-primed λ gt11 library. To isolate cDNA extending to the 5'-end of mRNA (clone R43-E), a rabbit skeletal muscle cDNA library was constructed using random primed cDNA with λ gt11 vector (Stratagene). All cDNA inserts were sequenced either on an Applied Biosystems Incorporated Automatic Sequencer or manually by the dideoxy chain termination method³¹. Sequences were analysed with the Genetics Computer Group (Wisconsin package) and PCGene (Intelligenetics) software.

FIG. 2 Expression of different regions of the 43–156K DAG precursor protein as glutathione S-transferase (GST) fusion proteins. *a*, Schematic structures of the GST fusion proteins. Each fusion protein consists of GST and the indicated region of 97K precursor protein. Fusion protein-A (FP-A) contains residues 665–856 corresponding to the cDNA R43-A found in the expression library with affinity-purified antiserum against 43K DAG. Fusion protein-C (FP-C) contains residues 857–895 which are the C-terminal 38 amino acids of the 97K precursor polypeptide, and does not overlap with FP-A. Fusion protein-B (FP-B) contains residues 367–863 and thus overlaps with FP-A and FP-C, and has a portion of the N-terminal region of the 97K precursor, which is not present in the mature 43K DAG. Fusion protein-D (FP-D) contains residues 62–438 and thus contains only the N-terminal region of the 97K precursor polypeptide. The M_r and amino-acid positions of each fusion protein are indicated. The 97K protein precursor sequence corresponding to the 156K DAG is shown as a hatched box whereas the region of the 43K DAG is shown as an open box. Predicted transmembrane and signal sequences are shown as a solid box. *b*, Immunoblot analysis of fusion proteins and the dystrophin-glycoprotein complex (DGC). Samples of partially purified fusion proteins were separated on 3–12% SDS polyacrylamide gels and transferred to nitrocellulose. Nitrocellulose transfers were stained with affinity-purified sheep polyclonal antibodies to the fusion protein constructs: FP-C (anti FP-C), FP-A (anti FP-A), FP-B (anti FP-B), FP-D (anti FP-D) or with a polyclonal rabbit antibody to the synthetic peptide representing the 15 amino acids of the carboxyl terminus of the 97K precursor protein (43 C-term). The M_r standards ($\times 10^{-3}$) are indicated on the left and the mature 43K and 156K components of DGC are indicated on the right (43 DAG, 156 DAG).

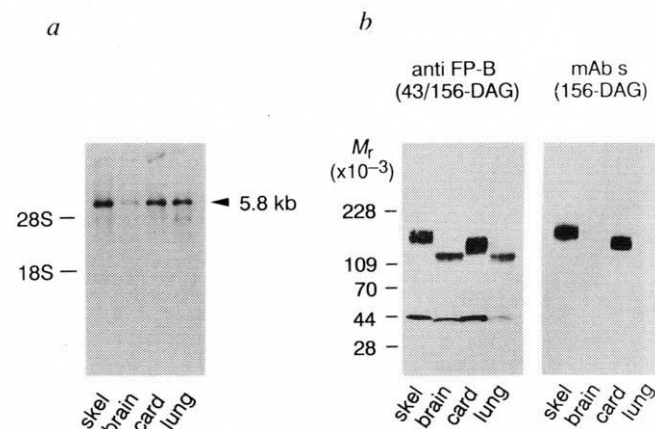
METHODS. A set of pGEX vectors³² were used to express various fragments of DNA for the 97K precursor protein as *E. coli* fusion proteins. The correct construction of the recombinant plasmids was verified by restriction mapping. To construct FP-A, the *EcoRI* insert from R43-A clone with the size of 0.6 kb was cloned into the *EcoRI* site of pGEX-1. FP-B was constructed by ligation of the *EcoRI* insert from R43-R clone (1.5 kb) into the *EcoRI* site of pGEX-2T. FP-C was made by ligation into the *BamHI* site of pGEX-1 the *BamHI* fragment of cDNA R43-C, containing C-terminal sequence with stop codon, representing the last 38 amino acids. For the FP-D construct, *EcoRI* insert (1.2 kb) from R43-D was inserted into pGEX-2T vector digested with *EcoRI*. Each recombinant molecule was introduced in *E. coli* DH5 α cells. Overnight cultures were diluted 1:10, incubated for 1 h and induced for 2 h with 1 mM IPTG. Cells were resuspended in PBS and sonicated. Fusion proteins



were purified from supernatant by affinity chromatography on glutathione-Sepharose (Pharmacia) and eluted with 5 mM glutathione. Dystrophin-glycoprotein complex was isolated as previously described⁹. Sheep polyclonal antibodies to the purified DGC were produced as previously described¹² and anti-fusion protein antibodies were affinity purified from polyclonal antiserum as previously described⁹. A peptide representing the 15 C-terminus amino acids of the 97K cDNA (PKNMTPYRSPPPYVP) was obtained from the HMMI Peptide Facility (Washington University, St. Louis) as the N-terminal *p*-benzylbenzoyl-peptide photoprobe³³. Peptide was conjugated to keyhole limpet haemocyanin, mixed with Freund's complete adjuvant and injected into a rabbit as described³⁴. SDS-PAGE³⁵ was carried out on 3–12% gradient gels in the presence of 1% 2-mercaptoethanol and transferred to nitrocellulose for immunoblot analysis as previously described³⁶.

FIG. 3 Tissue distribution of 43–156K DAG. *a*, Northern blot analysis is with 43–156K DAG-specific probe of mRNA from different tissues. Poly(A)⁺ RNA (4 μ g per lane) from rabbit skeletal muscle (skel), brain, cardiac muscle (card) and lung was transferred to nylon filters and hybridized with a ³²P-labelled insert from R43-A cDNA clone. The 43–156K DAG-specific mRNA with the size of 5.8 kb is indicated by the arrow. *b*, Immunostaining of identical nitrocellulose transfers of membranes from rabbit skeletal muscle, brain, cardiac muscle and lung stained with anti FP-B antibodies (43–156K specific) and a mixture of monoclonal antibodies VIA4₁ and IIH6 (156K DAG specific). M_r standards ($\times 10^{-3}$) are indicated on the left.

METHODS. Total RNA was isolated by homogenization in RNazol (Cinna/Bio-tex) followed by chloroform extraction. Poly(A)⁺ RNA was enriched by oligo-dT cellulose chromatography and resolved on 1.2% agarose gels containing 5% formaldehyde. RNA was transferred to Genescreen Nylon Membranes (NEN Research Products). Prehybridization was done at 42 °C in 5 $\times$ SSC, 5 $\times$ Denhardt's solution, 50% formamide, 10% dextran sulphate and 100 μ g ml⁻¹ of salmon-sperm DNA. Membranes were hybridized overnight at 42 °C at a specific activity of at least 1 $\times 10^6$ c.p.m. ml⁻¹. Membranes were washed at 62 °C in 2 $\times$ SSC, 0.1% SDS and were exposed to film (X-OMAT AR, Kodak) at -80 °C. Total membranes were prepared from tissues homogenized in 7.5 vols of homogenization buffer (20 mM sodium pyrophosphate, 20 mM sodium phosphate monohydrate, 1 mM MgCl₂, 0.3 M sucrose, 0.5 mM EDTA, pH 7.0) using a Polytron PTS-10-S probe (Kinematic GmbH, Luzern) in the presence of a protease inhibitor cocktail¹⁰. Homogenates were centrifuged for 15 min at 1,100g and the supernatant filtered



through cheesecloth. The supernatants of four repeated homogenizations were combined and centrifuged for 35 min at 140,000g. The final membrane preparations were KCl-washed as previously described¹⁰. Immunoblot analysis was done as described in the legend to Fig. 2 using 250 μ g of skeletal muscle membranes and 500 μ g of non-muscle membranes.

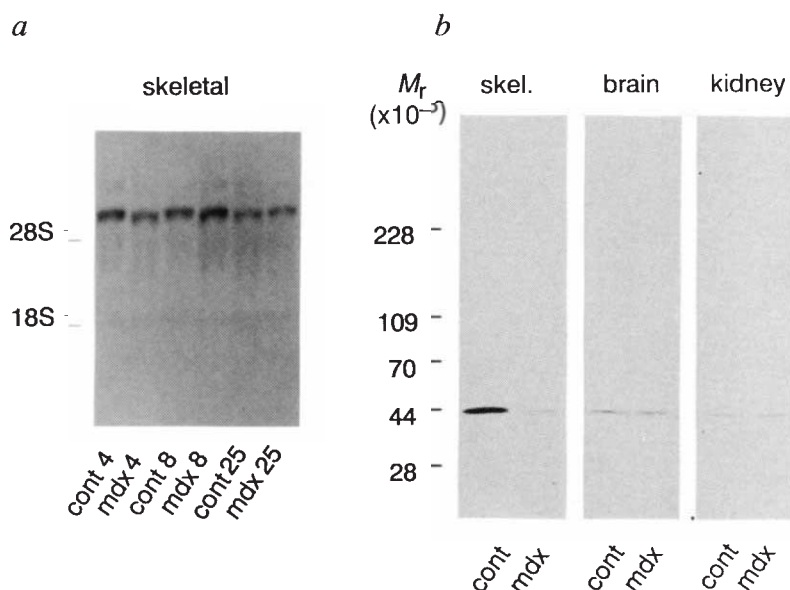


FIG. 4 Comparison of 43-156K DAG expression in muscle and non-muscle tissues of normal and mdx mice. *a*, Skeletal muscle RNA (20 µg per lane) of control mice of ages 4, 8 and 25 weeks (cont 4, cont 8, cont 25) and skeletal muscle RNA of mdx mice of ages 4, 8 and 25 weeks (mdx 4, mdx 8, mdx 25) hybridized with 32 P-labelled cDNA inset from R43-B clone. Immunoblot analyses were done as described in the legend to Fig. 3. *b*, Nitrocellulose transfers of skeletal muscle, brain and kidney membranes from control (cont) and mdx mice (mdx) were stained with affinity-purified anti-FP-A (43K DAG specific) antibodies. M_r standards are indicated. Total RNA and membranes from control and mdx mice were prepared as described in Fig. 3.

one site of *N*-linked glycosylation and numerous *O*-linked glycosylation sites. The deduced amino-acid sequence of the N-terminal portion of the precursor protein indicates a ~56K core peptide for 156K DAG. Thus, carbohydrate moieties seem to constitute up to two-thirds of the M_r of the 156K DAG which suggests that the 156K DAG may be a proteoglycan. The exact modifications involved in the processing of the N-terminal portion of the precursor polypeptide to the 156K DAG are not known, but because of the size difference between the 56K protein core and the mature 156K DAG they could involve the addition of glycosaminoglycan chains^{17,18}.

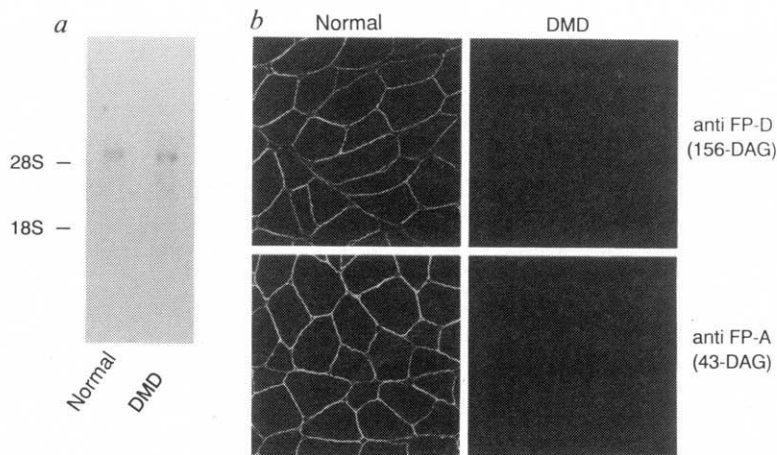
The predicted features of 156K DAG made it the most probable candidate for interaction with components in the extracellular space. We have demonstrated, using two different techniques, that extracellular 156K DAG binds laminin which is a major component of the extracellular matrix. A number of laminin-binding proteins have been previously identified in skeletal muscle¹⁹⁻²¹ but the 156K DAG does not seem to be related to any of these proteins. In addition, the sequence of the 43-156K DAG indicates that it is not related to integrins or cadherin. The exact function of laminin binding by the 156K DAG is not known. It is possible that the 156K DAG is involved in the attachment, spreading and growth of myoblasts on laminin substrate, which are independent of integrin²².

The broad tissue distribution of the 43-156K DAG precursor argues for an important role of the glycoproteins in membrane organization in different tissues and might indicate the existence

of the entire glycoprotein complex in non-muscle tissues. Absence of significant amounts of dystrophin in the examined non-muscle tissues suggests that in non-muscle tissues 43-156K DAG is involved in formation of a different type of complex where dystrophin may be replaced by another cytoskeleton component. The conditions used to identify the 156K DAG as a laminin-binding protein are similar to those used for the identification of crinin as a laminin-binding protein²³. The apparent M_r of crinin is also very similar to the protein we have identified in brain membranes with the '156K'-specific antibody.

How the deficiency of dystrophin causes muscle cell necrosis and ultimately leads to muscle weakness in DMD is not known. Our findings suggest that the function of dystrophin is to link the subsarcolemma membrane cytoskeleton through a trans-membrane complex to an extracellular glycoprotein which binds laminin. Because the absence of dystrophin leads to the loss of all the dystrophin-associated proteins^{8,12} our results suggest that dystrophin-deficient muscle fibres may lack the normal interaction between the sarcolemma and the extracellular matrix. The disruption of this linkage between the sarcolemma and the extracellular matrix may be responsible for the increased osmotic fragility of dystrophic muscle²⁴ or the alteration of specific Ca^{2+} regulatory mechanisms²⁵ either of which may lead to excessive influx of Ca^{2+} ions in dystrophic muscle²⁶. The histopathological analysis of DMD muscle^{27,28} supports this hypothesis because a breakdown of the sarcolemma membrane

FIG. 5 Characterization of 43-156K DAG in normal and DMD skeletal muscle. *a*, Northern blot analysis of total RNA from normal and DMD skeletal muscle probed with R43-B cDNA. Northern blot analysis with human normal and DMD skeletal muscle RNA was done as described in Fig. 3. *b*, Immunofluorescence analysis of transverse cryosections of normal and DMD skeletal muscle stained with affinity-purified sequence-specific antibodies against FP-D (156K DAG specific) or FP-A (43K DAG specific). Immunofluorescence microscopy of transverse cryosections from normal and DMD muscle was done as previously described^{10,12}.



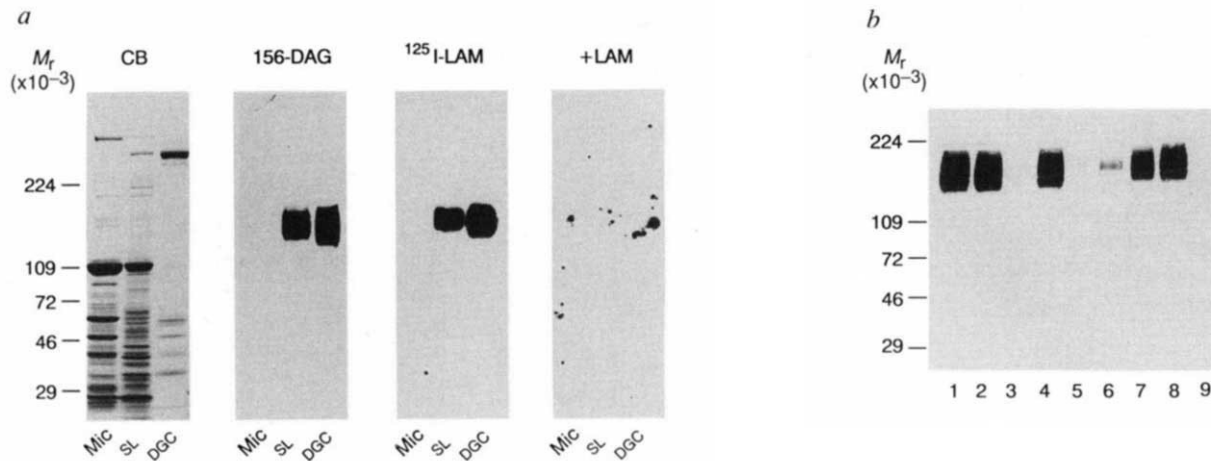


FIG. 6 Laminin binding to the 156K DAG. *a*, Coomassie blue-stained SDS polyacrylamide gel (CB), corresponding immunoblot stained with monoclonal antibody IIH6 to 156K DAG or autoradiograms of identical nitrocellulose transfers of electrophoretically separated proteins of crude rabbit skeletal muscle membranes (Mic), sarcolemma membranes (SL) or dystrophin-glycoprotein complex (DGC) incubated with $0.1 \mu\text{g ml}^{-1}$ ^{125}I -labelled laminin ($\sim 1.7 \mu\text{Ci } \mu\text{g}^{-1}$) in the absence (^{125}I -LAM) or presence of a 1,000-fold excess of unlabelled laminin (+LAM). *b*, Coimmunoprecipitation of 156K DAG using anti-laminin antibody. Shown is the relative amount of 156K DAG in pH 12 extracts of crude rabbit surface membranes (lane 1), in anti-laminin/protein A-Sepharose voids (lanes 2 and 6), associated with anti-laminin/Protein A-Sepharose (lanes 3 and 7), protein A-Sepharose voids (lanes 4 and 8) or associated with protein A-Sepharose (lanes 5 and 9) in the absence (lanes 2-5) or presence (lanes 6-9) of $50 \mu\text{g}$ laminin. M_r standards ($\times 10^{-3}$) are indicated on the left of each panel.

METHODS. Rabbit skeletal muscle crude membranes, sarcolemma membranes and dystrophin-glycoprotein complex were electrophoretically separated on 3–12% SDS polyacrylamide gels in the presence of 1% 2-mercaptoethanol and transferred to nitrocellulose (see legend to Fig. 2).

Purified laminin mouse EHS (Sigma) was iodinated with ^{125}I NaI using a lactoperoxidase/glucose oxidase reaction by the Diabetes, Endocrinology Research Center at the University of Iowa. The ^{125}I -labelled laminin overlay procedure of Smalheiser and Schwartz²³ was done as described except that nitrocellulose transfers were blocked with 5% non-fat dry milk in 150 mM NaCl, 50 mM sodium phosphate, pH 7.5. To test for coimmunoprecipitation of laminin and the 156K dystrophin-associated glycoprotein, pH 12 extracts⁸ of rabbit skeletal muscle surface membranes were incubated for 24 h at 4°C with gentle mixing in buffer A (0.14 M NaCl, 1 mM CaCl_2 , 1 mM MgCl_2 , 10 mM triethanolamine-HCl, pH 7.6) in the absence or presence of $50 \mu\text{g}$ purified mouse EHS laminin (Sigma), then incubated for an additional 24 h at 4°C with $100 \mu\text{l}$ of either protein A-Sepharose or anti-laminin/protein A-Sepharose which had been equilibrated with 3% BSA in buffer A and washed four times with buffer A. The Sepharose was pelleted by a brief centrifugation, the supernatant (void) decanted and the Sepharose washed three times with buffer A. Equivalent volumes of the resulting voids and washed Sepharose pellets were analysed by SDS-polyacrylamide gel electrophoresis and immunoblotting using monoclonal antibody IIH6 which is specific for the 156K DAG.

precedes muscle cell necrosis and the basal lamina seems to separate from the sarcolemma in early steps of DMD. In addition, because the extracellular matrix of the adult tissue is a scaffold that is required to allow repair after injury, the absence of the interaction between the sarcolemma and the extracellular matrix may also render dystrophic muscle fibres more prone to injury and less able to repair injury.

The normal production of the mRNA for the 43–156K DAG in dystrophic muscle is important for potential DMD therapies. It was unclear how the absence of dystrophin leads to the loss of the dystrophin-associated glycoproteins but these results indicate that dystrophin-associated glycoproteins are produced in

dystrophic muscle. But in the absence of dystrophin the dystrophin-associated glycoproteins may not be properly assembled and/or integrated into the sarcolemma or may be degraded. Our results suggest that restoring dystrophin by myoblast transfer or gene therapy may stabilize and restore normal dystrophin-associated glycoprotein levels in DMD muscle. These results provide strong evidence for the structural role of dystrophin in muscle integrity, demonstrate that the 43–156K DAG (dystroglycan) is a new laminin-binding glycoprotein and suggest that the function of dystrophin-glycoprotein complex is to provide linkage between the sarcolemma and the extracellular matrix. □

Received 21 November 1991; accepted 16 January 1992.

- Hoffman, E. P., Brown, R. H. & Kunkel, L. M. *Cell* **51**, 919–928 (1987).
- Koenig, M., Monaco, A. P. & Kunkel, L. M. *Cell* **53**, 219–228 (1988).
- Arahata, K. *et al. Nature* **333**, 861–866 (1988).
- Hoffman, E. P. *et al. N. Engl. J. Med.* **318**, 1363–1368 (1988).
- Watkins, S. C., Hoffman, E. P., Slayter, H. S. & Kunkel, L. M. *Nature* **333**, 863–866 (1988).
- Cooper, B. J. *et al. Nature* **334**, 154–156 (1988).
- Campbell, K. P. & Kahl, S. D. *Nature* **338**, 259–262 (1989).
- Ervasti, J. M., Ohlendieck, K., Kahl, S. D., Gaver, M. G. & Campbell, K. P. *Nature* **345**, 315–319 (1990).
- Ervasti, J. M. & Campbell, K. P. *Cell* **66**, 1121–1131 (1991).
- Ohlendieck, K., Ervasti, J. M., Snook, J. B. & Campbell, K. P. *J. Cell Biol.* **112**, 135–148 (1991).
- Ohlendieck, K. & Campbell, K. P. *FEBS Lett.* **283**, 230–234 (1991).
- Ohlendieck, K. & Campbell, K. P. *J. Cell Biol.* **115**, 1685–1694 (1991).
- Devi, L. *FEBS Lett.* **280**, 189–194 (1991).
- Ehrig, K., Leivo, I., Argraves, W. S., Ruoslahti, E. & Engvall, E. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3264–3268 (1990).
- Hunter, D. D., Shah, V., Merlie, J. P. & Sanes, J. R. *Nature* **338**, 229–234 (1989).
- Sanes, J. R., Engvall, E., Butkowsky, R. & Hunter, D. D. *J. Cell Biol.* **111**, 1685–1699 (1990).
- Hassell, J. R., Kimura, J. H. & Hassell, V. C. *A. Rev. Biochem.* **55**, 539–567 (1986).
- Höök, M., Kjellén, L., Johansson, S. & Robinson, J. A. *Rev. Biochem.* **53**, 847–869 (1984).
- Lesot, H., Kühn, U. & von der Mark, K. *EMBO J.* **2**, 861–865 (1983).
- Clegg, D. O., Helder, J. C., Hann, B. C., Hall, D. E. & Reichardt, L. F. *J. Cell Biol.* **107**, 699–7095 (1988).
- Mecham, R. P. *FASEB J.* **5**, 2538–2546 (1991).

- Volk, T., Fessler, L. I. & Fessler, J. H. *Cell* **63**, 525–536 (1990).
- Smalheiser, N. R. & Schwartz, N. B. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6457–6461 (1987).
- Merke, A. & Jockusch, M. *Nature* **349**, 69–71 (1991).
- Franco, A. & Lansman, J. B. *Nature* **344**, 670–673 (1990).
- Turner, P. R., Fong, P., Denetclaw, W. F. & Steinhardt, R. A. *J. Cell Biol.* **115**, 1701–1712 (1991).
- Bonilla, E. & Moggio, E. *Neurology* **36**, 171 (1986).
- Carpenter, S. & Karpatis, G. *Brain* **102**, 147–161 (1979).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
- Jay, S. D. *et al. Science* **248**, 490–492 (1990).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Smith, D. B. & Johnson, K. S. *Gene* **67**, 31–40 (1988).
- Gorka, J., McCourt, D. W. & Schwartz, B. D. *Peptide Res.* **2**, 1–5 (1989).
- McPherson, P. S. *et al. Neuron* **7**, 17–25 (1991).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Campbell, K. P. *et al. J. biol. Chem.* **262**, 6460–6463 (1987).

ACKNOWLEDGEMENTS. We thank S. D. Kahl, D. Zoloto, J. Hsu and C. R. Moomaw for technical assistance, D. S. Lange and J.-C. Fan for secretarial assistance, S. Numa for rabbit skeletal muscle cDNA, E. Engvall and J. Sanes for antibodies to merosin and S-laminin, respectively and J. Gorka, Howard Hughes Medical Institute, K. P. C. is an investigator of the Howard Hughes Medical Institute; J.M.E. is the Carl M. Pearson Fellow of the Muscular Dystrophy Association. This work was also supported by the Muscular Dystrophy Association. The nucleotide sequence of the 43–156K DAG will appear in GENBANK accession number X64393.

Infrared helium emission lines from Cygnus X-3 suggesting a Wolf-Rayet star companion

M. H. van Kerkwijk^{*†‡}, P. A. Charles^{§||}, T. R. Geballe[¶],
D. L. King[§], G. K. Miley[#], L. A. Molnar^{‡**},
E. P. J. van den Heuvel^{*†‡}, M. van der Klis^{*†}
& J. van Paradijs^{*†}

^{*} Astronomical Institute "Anton Pannekoek", University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands

[†] Center for High-Energy Astrophysics, Kruislaan 409, 1098 SL Amsterdam, The Netherlands

[‡] Institute for Theoretical Physics, University of California at Santa Barbara, Santa Barbara, California 93106, USA

[§] Royal Greenwich Observatory, Apartado 321,

38780 Santa Cruz de La Palma, Tenerife, Canary Islands

^{||} Department of Astrophysics, Nuclear Physics Building, Keble Road, Oxford OX1 3RH, UK

[¶] Joint Astronomy Centre, 665 Komohana Street, Hilo, Hawaii 96720, USA

[#] Sterrewacht, Niels Bohrweg 2, PO Box 9513, 2300 RA Leiden, The Netherlands

^{**} Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa 52242, USA

CYGNUS X-3 is one of the most luminous X-ray sources in the Galaxy^{1,2}, a bright infrared source³ and a radio source that undergoes huge outbursts⁴. The system is a binary, presumably a neutron star plus companion, with a 4.79-h orbital period that modulates the X-ray and infrared emission^{5,6} and that increases on a 600,000-year timescale^{7,8}. Radio observations reveal the presence of a relativistic jet⁹. The nature of Cyg X-3 has remained unclear, however, in part because the large interstellar extinction³ in its direction prevents optical spectroscopy. Upper limits on spectral features in the near infrared have been reported previously¹⁰, but only with recent instrumental improvements have we become able to identify spectral features in the near infrared I and K bands. These are found to be characteristic of Wolf-Rayet stars: strong, broad emission lines of He I and He II, but no strong hydrogen lines. These observations strongly suggest the presence of a dense wind in the Cyg X-3 system, and may indicate that the companion is a fairly massive helium star, as had been predicted¹¹ by a model in which the present system is a descendant of a massive X-ray binary.

The I-band spectrum (Fig. 1) was obtained in service time on 22 June 1991, with ISIS on the 4.2-m William Herschel Telescope on La Palma. The K-band spectrum (Fig. 2) was taken with the

TABLE 1 Characteristics of the observed features

λ (μm)	Identification	-EW ($\text{\AA}$)	FWHM* (10^3 km s^{-1})
1.0120 (4)	He II (5-4)	90 (20)	2.7 (3)
2.060 (1)	He I $2p^1P^o - 2s^1S$	50 (10)	2.1 (3)
2.112 (1)	He I $4s^3S - 3p^3P^o$	20 (6)	2.4 (3)
	He I $4s^1S - 3p^1P^o$		
	He II (14-8)		
2.166 (1)	He I (7-4)	26 (3)	2.9 (7)
	H I (7-4)		
2.188 (1)	He II (10-7)	29 (4)	2.9 (7)
2.346 (2)	He II (13-8)	12 (2)	2.6 (4)

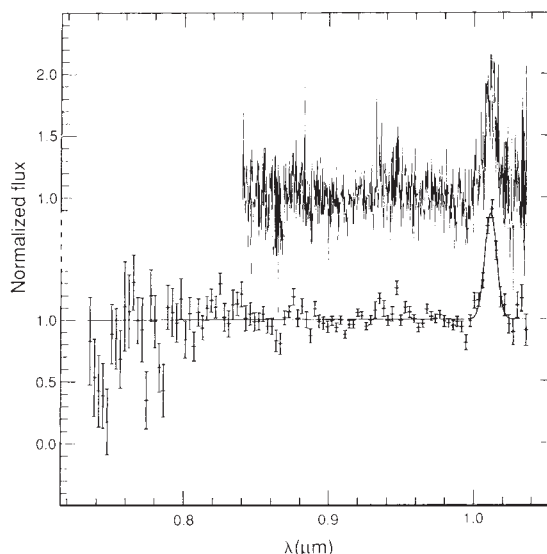
Errors indicate 90% confidence levels. EW is equivalent width; FWHM is full width at half maximum.

* As observed; not corrected for the instrumental resolution.

CGS4 on the United Kingdom 3.8-m Infrared Telescope (UKIRT) on Hawaii, during service observations on 29 June 1991. The spectra show a number of strong, very broad emission features. For each line the identification, equivalent width and full width at half-maximum are listed in Table 1. All the lines can be identified as either He I or He II lines, except the line at 2.166 μm . Given the strength of the He II line at 2.347 μm (13-8), the 2.166- μm line cannot be due solely to He II (14-8). Other possible contributors are H I (7-4) ($\text{Br}\gamma$) and several lines of the He I (7-4) transition array.

The strong, broad emission lines of both He I and He II, and the lack of strong hydrogen emission, are reminiscent of a Wolf-Rayet (WR) spectrum (for a definition, see for example refs 12, 13). We have therefore compared our spectra with infrared spectra of WR stars¹⁴⁻¹⁹. The absence of carbon emission lines such as C III at 0.971 μm and C IV at 2.075 μm indicates that the spectra are not of the carbon-rich WC spectral sequence^{14,19}. For the nitrogen-rich WN sequence, however, all expected lines are observed¹⁴⁻¹⁸. (Possible nitrogen lines are expected to be weak. If present, they are blended with the He I line at 2.112 μm .) Both the relative and absolute strengths of He I at 2.112 μm and He II at 2.188 μm indicate spectral subtype WN7. The equivalent widths of the other He II lines are consistent with this classification, but the He I line at 2.058 μm is stronger than in any of the eight WN spectra we have available for comparison¹⁵⁻¹⁸. The blend at 2.166 μm is somewhat weaker than in the two available WN7 spectra^{16,18}, possibly indicating a lower hydrogen abundance in Cyg X-3. The widths of the lines are similar to those observed in WN stars^{12,14-18}.

FIG. 1 The I-band spectrum obtained on 22 June 1991, 4:45 UT. The reduced, normalized spectrum, with the noisy short-wavelength portion omitted, is shown at the top. The resolution is $\sim 200 \text{ km s}^{-1}$. The emission feature at 1.012 μm is He II (5-4). The full spectrum, averaged over 11 pixels, is at the bottom, with a gaussian fit to the line superposed. In our data reduction, we used the spectrum of the nearby star A (see finding chart in ref. 29; the slit had been positioned on stars A and C to be certain that Cyg X-3 was observed), with its stellar features (Paschen and Ca II absorption) artificially removed, to correct for the telluric features in the Cyg X-3 spectrum. As a plot of $\log F_\lambda$ against $\log \lambda$ for the ratioed continuum is linear, the spectrum was divided by a power-law fit to produce the result shown here.



The presence of lines of relatively low excitation in the infrared spectrum seems inconsistent with any model for Cyg X-3 that requires a very high temperature ($>10^6$ K) in the infrared-flux-producing region (such as an accretion disc corona model²⁰, or, more generally, any model in which the size of the infrared object is comparable to, or smaller than, the orbital separation⁵). Instead, the WR spectrum probably indicates¹³ that a strong wind is present in Cyg X-3, which in the infrared is optically thick out to radii much larger than the orbital separation. Such a wind might be due to X-ray irradiation of either the accretion disc or a low-mass companion. Alternatively, it could be intrinsic to the companion: that is, the companion could be a WR star. For lack of predictions, the irradiation models are currently difficult to test. Below, we therefore confine ourselves to the WR model.

If the companion is a WR star, then given the absolute K magnitude of Cyg X-3 of <-5 (ref. 3), it can only be a 'classical' WR star²¹: the remnant helium core of a massive O-type or early B-type star that has lost (most of) its hydrogen-rich envelope, and is now in the stage of helium burning or beyond. Hence, the system would be composed of a compact object and a helium star. This was proposed earlier by Van den Heuvel and De Loore¹¹, who predicted the existence of such systems from considerations of the evolutionary fate of massive X-ray binaries such as Cen X-3 and SS 433. In these binaries the primary is a massive, early-type star, which expands in the course of its evolution, and eventually overflows its critical surface (Roche lobe). The resulting mass-transfer to the secondary is unstable, and causes the orbit to shrink on a thermal timescale. If the original orbit was wide enough the system can stabilize, but only after almost the whole hydrogen-rich envelope of the primary has been expelled. The remaining helium star is likely finally to explode as a supernova. If the system is not disrupted, a binary such as the Hulse-Taylor pulsar PSR1913+16, composed of two neutron stars in a highly eccentric orbit, could be formed²².

Stellar model calculations^{23,24} show that a helium star in Cyg

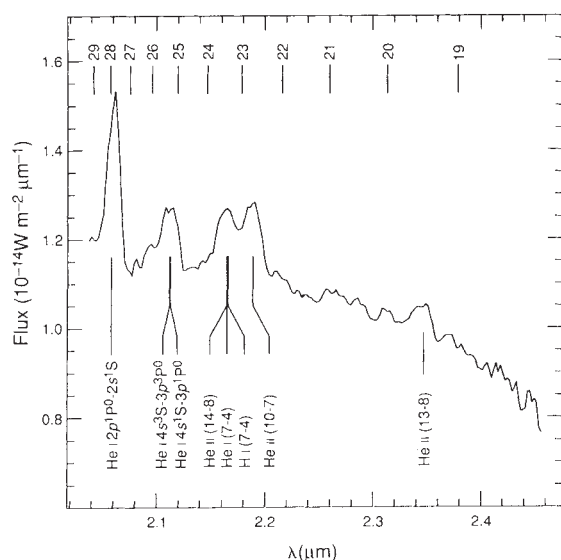


FIG. 2 The K-band spectrum obtained on 29 June, 1991, 15:00 UT. Shown is the reduced spectrum, corrected for telluric absorption and flux-calibrated using a spectrum of HR7796 (with its H β Br γ absorption line artificially removed). The resolution is $\sim 1,500$ km s $^{-1}$. Below the spectrum, the identifications of the stronger lines are given. Above the spectrum, the rest wavelengths of the He II ($n-9$) lines are shown, as several of these coincide with marginally significant emission features in the spectrum. A possibly real feature is also seen at $2.286 \mu\text{m}$. For this line no likely identification has been found. (It may have the same origin as the unidentified feature at $2.287 \mu\text{m}$ observed in spectra of planetary nebulae³⁰.)

X-3 would fit well inside its Roche lobe: the radius of a 1, 5, 10, 20 or $60 M_{\odot}$ helium star is 0.2, 0.6, 0.9, 1.4 or $2.6 R_{\odot}$, respectively²⁴, whereas the radius of the Roche lobe would be 0.6, 1.3, 1.7, 2.3 or $3.8 R_{\odot}$, respectively (for a $1.4 M_{\odot}$ compact object, the canonical mass of a neutron star). On the other hand, the 'zero-velocity' radii derived from model fitting to WR spectra²⁵, which should also be a measure of helium star radii, are never smaller than $2-3 R_{\odot}$, and in most cases range from 5 to $20 R_{\odot}$. The cause of this discrepancy is unknown. It may indicate that the velocity law assumed for the WR atmosphere, which is used to extrapolate from the radius at which the continuum and the emission lines are observed to the 'zero-velocity' radius, is not correct.

The observed rate of increase of the orbital period of Cyg X-3 (ref. 8) allows one to estimate the rate of stellar-wind mass loss from the companion. For spherically symmetric mass loss one obtains²⁶

$$\frac{\dot{M}}{M} = \frac{\dot{P}}{2P} = 0.8 \times 10^{-6} \text{ yr}^{-1}$$

where $\dot{M}$ is mass-loss rate, and M the total mass of the system. For a $1.4 M_{\odot}$ compact object and a $10 M_{\odot}$ WR star, we find $\dot{M} \approx 0.9 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$, well within the range $(0.8-8) \times 10^{-5} M_{\odot} \text{ yr}^{-1}$ found for WR stars¹².

The mass-loss rate can be used to check whether the column density in the WR wind between the compact object and the observer is such that the observed modulation of the X-ray flux along the orbit can be obtained. Previous modelling has shown^{27,28} that this requires the electron-scattering optical depth τ_{es} to be close to unity. A first-order estimate of τ_{es} can be made under the assumption of a spherically symmetric pure helium wind that flows at a constant velocity v_w . It is easy to show that at ascending and descending node (both stars in the plane of the sky)

$$\tau_{\text{es}} \approx 0.35 \gamma \left(\frac{\dot{M}}{10^{-5} M_{\odot} \text{ yr}^{-1}} \right) \left(\frac{M}{11.4 M_{\odot}} \right)^{-1/3} \times \left(\frac{v_w}{1,000 \text{ km s}^{-1}} \right)^{-1}$$

where γ is the number of electrons per helium ion. The main contribution to τ_{es} comes from the region close to the compact object, in the inner part of the wind. In this region the wind will be almost completely ionized ($\gamma = 2$), and its velocity a few tenths of the terminal velocity of $\sim 2,000$ km s $^{-1}$ (as indicated by the widths of the lines; in WN stars $1,500-4,000$ km s $^{-1}$ is observed¹²). Hence, τ_{es} will be about unity, as required.

We conclude that the infrared spectra of Cygnus X-3 presented here indicate that a strong wind is present in the system, which in the infrared is optically thick up to well outside the system. We have shown that such a wind may originate from a WR companion. If this is the case, the system is in a later evolutionary stage of massive X-ray binaries¹¹. $\square$

Received 27 August; accepted 6 December 1991.

- Giacconi, R., Gorenstein, P., Gursky, H. & Waters, J. R. *Astrophys. J.* **148**, L119-L127 (1967).
- Dickey, J. M. *Astrophys. J.* **273**, L71-L73 (1983).
- Becklin, E. E., Kristian, J., Neugebauer, G. & Wynn-Williams, C. G. *Nature phys. Sci.* **239**, 130-131 (1972).
- Gregory, P. C. *et al. Nature phys. Sci.* **239**, 114-117 (1972).
- Becklin, E. E. *et al. Nature* **245**, 302-304 (1973).
- Mason, K. O. *et al. Astrophys. J.* **207**, 78-87 (1976).
- Manzo, G., Molteni, D. & Robba, N. R. *Astr. Astrophys.* **70**, 317-318 (1978).
- Van der Klis, M. & Bonnet-Bidaud, J. M. *Astr. Astrophys.* **214**, 203-208 (1989).
- Molnar, L. A., Reid, M. J. & Grindlay, J. E. *Astrophys. J.* **331**, 494-508 (1988).
- Katz, J. I., Wright, E. L. & Lawrence, C. R. *Astr. J.* **89**, 1604-1605 (1984).
- Van den Heuvel, E. P. J. & De Loore, C. *Astr. Astrophys.* **25**, 387-395 (1973).
- Conti, P. S. & Underhill, A. B. *O-type Stars and Wolf-Rayet Stars* (NASA SP-497, 1988).
- Conti, P. S. in *IAU Symp. No. 99: Wolf-Rayet Stars: Observations, Physics, Evolution* (eds de Loore, C. W. H. & Willis, A. J.) 3-19 (Kluwer, Dordrecht, 1990).
- Conti, P. S., Massey, P. & Vreux, J. M. *Astrophys. J.* **354**, 359-371 (1990).
- Williams, P. M. in *IAU Symp. No. 99: Wolf-Rayet Stars: Observations, Physics, Evolution* (eds de Loore, C. W. H. & Willis, A. J.) 73-78 (Kluwer, Dordrecht, 1990).
- Hillier, D. J. in *IAU Symp. No. 99: Wolf-Rayet Stars: Observations, Physics, Evolution* (eds de Loore, C. W. H. & Willis, A. J.) 225-228 (Kluwer, Dordrecht, 1990).

17. Hillier, D. J., Jones, T. J. & Hyland, A. R. *Astrophys. J.* **271**, 221–236 (1983).
18. Hillier, D. J. *Astr. J.* **90**, 1514–1518 (1985).
19. Smith, L. F. & Hummer, D. G. *Mon. Not. R. astr. Soc.* **230**, 511–534 (1988).
20. White, N. E. & Holt, S. S. *Astrophys. J.* **257**, 318–337 (1982).
21. Van der Hucht, K. A., Conti, P. S., Lundström, I. & Stenholm, B. *Space Sci. Rev.* **28**, 227–306 (1981).
22. Flannery, B. P. & Van den Heuvel, E. P. J. *Astr. Astrophys.* **39**, 61–67 (1975).
23. Divine, N. *Astrophys. J.* **142**, 824–840 (1965).
24. Langer, N. *Astr. Astrophys.* **210**, 93–113 (1989).
25. Schmutz, W., Hamann, W.-R. & Wessolowski, U. *Astr. Astrophys.* **210**, 236–248 (1989).
26. Davidsen, A. & Ostriker, J. P. *Astrophys. J.* **189**, 331–338 (1974).
27. Pringle, J. E. *Nature* **247**, 21–22 (1974).
28. Bonnet-Bidaud, J. M. & Van der Klis, M. *Astr. Astrophys.* **101**, 299–304 (1981).
29. Wagner, R. M., Kreidl, T. J., Bus, S. J. & Williams, W. *Astrophys. J.* **346**, 971–975 (1989).
30. Geballe, T. R., Burton, M. G. & Isaacman, R. *Mon. Not. R. astr. Soc.* **253**, 75–79 (1991).

ACKNOWLEDGEMENTS. We thank P. S. Conti for his constructive review. The William Herschel Telescope is operated by the Royal Greenwich Observatory in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. UKIRT is operated by the Royal Observatory, Edinburgh, on behalf of the SERC. This research was supported in part by the Netherlands Organisation for Scientific Research NWO (M.H.v.K., M.v.d.K.), the Leids Kerkhoven-Bosscha Fonds (M.H.v.K.) and the NSF (M.H.v.K., L.A.M., E.P.J.v.d.H.).

Three-dimensional motion in the radio jet of the binary system R Aquarii

Harry J. Lehto*† & Dean R. H. Johnson‡

* Department of Physics, University of Southampton, Highfield, Southampton SO9 5NH, UK

‡ University of Wales, Cardiff, UK

R AQUARI is a symbiotic binary system surrounded by a complex extended optical nebulosity¹. At radio and optical wavelengths a jet is seen to emerge from the central binary system^{2,3}. We have observed R Aqr using the Very Large Array. Comparison with earlier radio observations shows that five out of six bright components in the radio jet have moved. One radio component has the same proper motion as the optical Mira, the primary star of the binary. At a distance of 200 pc (refs 1, 4), the proper motions of the other components correspond to a tangential velocity of 44 to 160 km s⁻¹ with respect to the Mira. By combining these measurements with radial velocity determinations, we obtain a true three-dimensional velocity map of the radio jet, provided only that the observed proper motions indeed correspond to physical motions of emitting material. Our results rule out the possibility that the radio components in the jet were formed in a single explosive event, and suggest instead that they are 'bullets' ejected at ~20-yr intervals into a narrow cone. Alternatively, if the components move along the jet and are accelerated during the whole of their passage through the inner 7 arcsec (1,400 AU) of the system, ejection at ~40-yr intervals would lead to the disposition observed at present.

Symbiotic systems are long-period binaries consisting of a cool giant primary and a hot secondary star. R Aquarii is a unique symbiotic system because it contains as a primary a Mira variable with a period of 387 days. From its variability a distance of 250 pc has been estimated⁴. Kinematic studies of the outer and inner optical nebulosities¹ yield an estimated distance of 180 pc. Here we adopt a distance of 200 pc.

The radio continuum flux density ~8 mJy was measured⁵⁻⁷ in the late 1970s. A core-jet structure detected at optical wavelengths^{8,9} was also apparent in the first Very Large Array (VLA) measurements^{2,3}. Later VLA observations¹⁰⁻¹² showed that the system consists of five components (A', C1, C2, A, B). In 1986 a new jet component, D, was revealed^{13,14} by narrow-band H α imaging. The Mira is also a SiO maser source in radio wavelengths¹⁵⁻¹⁷, implying that the system is losing mass.

We observed R Aqr on 18 August 1991 for 196 minutes and 198 minutes at wavelengths of $\lambda = 3.6$ and 6.2 cm, respectively, with 19 operational antennas of the radio telescope in its largest,

A configuration. The data were edited and calibrated using the program AIPS at the National Radio Astronomy Observatory Arrays Operation Center in Socorro, New Mexico. The CLEANed 3-cm and 6-cm maps had a r.m.s. noise level of $\sigma = 31 \mu\text{Jy}$ and $33 \mu\text{Jy}$, respectively.

The 3-cm map shows four obvious components which we identified as A', C1, C2 and A. In addition to these, the 6-cm map shows structure corresponding to components B and D. We have thus detected component D at radio wavelengths. Components A', C1, C2, and A have sizes smaller than the 3-cm beam. Components C1 and A seem to be extended along position angles (PA) ~50° and ~14°, respectively. Components B and D are very extended and mostly resolved out in the untapered 3-cm map.

A comparison with earlier observations¹⁰⁻¹² showed that the positions of five of the six radio components had changed. We calculated for each component its proper motion, that is, its apparent motion perpendicular to the line of sight (Tables 1 and 2). The components clearly do not have a common proper motion. Component A' has a position close to that of the previous epoch. From the marginal proper motion alone, we cannot judge whether this component is a background source or a component of a 'counter-jet'. But the similarity of its radio spectrum to C2 and A, and the presence of a counter jet in optical¹⁸ and ultraviolet observations¹⁹, lead us to believe that it is associated with the R Aqr jet system. Details of the radio spectra of all the components will be discussed elsewhere (D.R.H.J. and H.J.L., manuscript in preparation). Component C1 has the smallest proper motion of the other five components, and because this is the same as the optical proper motion of the Mira star^{20,21} we believe that it represents the Mira, and that it is close to the base of the jet.

A previous model for the optical nebulosity¹ suggests that the inner optical nebulosity and the radio structure were created in a single explosion. The positions of the components extrapolated to the past should coincide at the location and time of the explosion. No such unique location or time is compatible with our measurements of proper motion.

The model of Kafatos *et al.*²² explains the radio and optical components as being ejected at periastron passage of the binary, at intervals of ~44 yr. In this model, radiation pressure from the central hot star on the dust in the components accelerates them to a terminal velocity within the first $\leq 30 \text{ AU} \approx 0.15''$. If radiation pressure is important, then the different proper motions that we observe imply either that the dust properties vary from one component to another, or that the column density of the dust varies, being highest in the components with highest relative proper motions, B and D. As the proper motions of the central star and the components were not available when this model²² was proposed, it is not surprising that the details of their geometry are not in agreement with our observations.

Assuming that each component had stayed at a constant velocity for most of its lifetime, we solved simultaneously from the observed positions of C1, C2, A, B and D for their respective proper motion vectors, times of zero separation, and the present position (α_0 , δ_0) of the core C1. The times of zero separation from this fit are 20.3, 40.0, 63.4 and 56.7 yr ago. The best solution for the proper motion of the core (+0.0029, -0.031) is slightly discordant with the optical proper motion. The best solutions for absolute proper motions for components C2, A, B and D are 0.076, 0.115, 0.129, 0.158'' yr⁻¹, in good agreement with the observed values. The calculated position angles of the absolute proper motion, which are the position angles of the ejection of the components, are +98°, +74°, +58° and +41° for components C2, A, B and D. If the components are ejected near the periastron of an eccentric orbit, then the orbital period of the system should be ~20 yr, or about half of the claimed period of 44 yr (ref. 23). Excess mass transfer and the ejection of the component could occur when the Mira is attempting to reach its maximum size in its 387-day cycle, causing strong Roche-lobe overflow. It

† Present address: Turku University Observatory, Tuorla, Piikkiö, Finland.

TABLE 1 Positions of the components in the jet of R Aqr

	New RA (1950.0) 23 h 41 min	New Dec (1950.0) −15° 33'	Old RA (1950.0) 23 h 41 m	Old Dec (1950.0) −15° 33'	Old epoch	New PA (°)	Distance from C1 (")
A'	14.205(3) s	44.08(8)"	14.193(8)	44.10(17)"	1987.73	−130.9(2.7)	1.43
C1	14.280(1) s	43.14(5)"	14.266(3)	43.00(5)"	1985.03	—	—
C2	14.327(2) s	42.68(7)"	14.293(3)	42.65(5)"	1985.03	55.9(4.2)	0.82
A	14.469(4) s	40.64(11)"	14.398(7)	40.85(10)"	1982.74	47.5(1.4)	3.70
B	14.599(8) s	37.13(15)"	14.493(7)	37.23(10)"	1982.74	37.5(1.0)	7.57
D	14.530(8) s	34.68(15)"	14.482(10)	35.13(15)"	1986.91	23.1(0.8)	9.20

The estimated 1σ accuracy of the fit is shown in parentheses. The epoch of our observation is 1991.63. The positions of components C1, C2 and A' were obtained by fitting simultaneously three gaussians to the combined image of these components in the 3-cm map. The position of A was obtained by fitting a single gaussian to the component in the 3-cm map. The positions of B and D are from single gaussian fits to the 6-cm maps. As a check we also determined the positions of A', C1, C2 and A from the 6-cm map. They were all within the 1σ estimated error of the 3-cm positions. The errors in the fits to B and D are larger than to other components because lower-resolution maps were used, and because of the extended structure of those components. The old epoch position of A' was obtained by re-analysing the data of ref. 27. The position of old epoch component D is from optical H α imaging. All position angles are given with respect to C1. The position of the Mira from optical observations^{20,21} corrected for proper motion is RA (1950.0) 23 h 41 min 14.267(6) s, Dec (1950.0) −15° 33' 42.62(16)". This is within 3σ of the position of C1.

seems that B and D may have been initially a single component which later broke apart.

The structure of the inner and outer nebulosities suggest that the axis of the disk (and the binary) is on PA $\sim 0^\circ$. This implies that the ejection of the components should have occurred in a nonpolar direction, possibly in the direction of L₃, the lagrangian point that is behind the hot component as seen from the Mira. From one orbit to the next, the maximum radius of the Mira would occur at a slightly different location on the orbit, causing a small difference in the ejection angles and hence in the observed position angles of the proper motion vectors of the components. If the ejection is towards the polar direction

of the accretion disk and not in its plane, then the 24° difference in the position angle of the velocity vector from one period to the next could be interpreted as being due to the precession of the disk.

An alternative model is that of a continuous jet, where the components move along the jet. This is suggested by the apparent connection between components A and C2 in our 6-cm map, and also, at optical wavelengths, by recent observations²⁴ with the Hubble Space Telescope. It seems that in this jet the magnitude of the relative proper motion increases with increasing distance from the Mira. Components C2, A, B and D follow a law $\mu \approx 0.0556\sqrt{d}$, where μ is the observed proper motion in

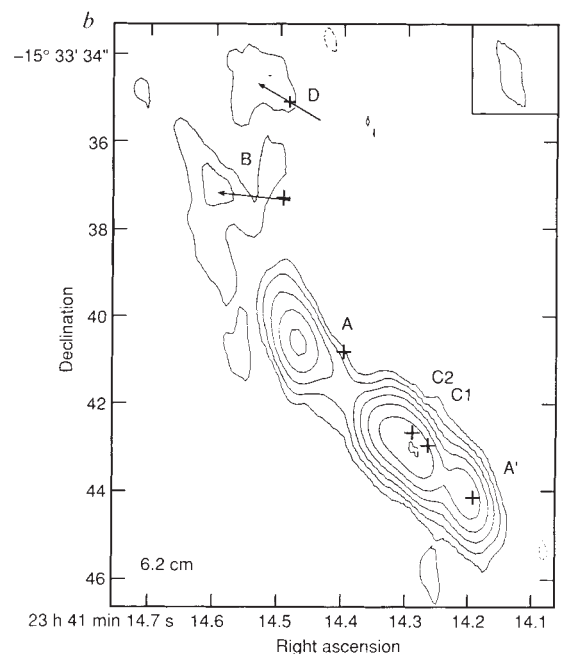
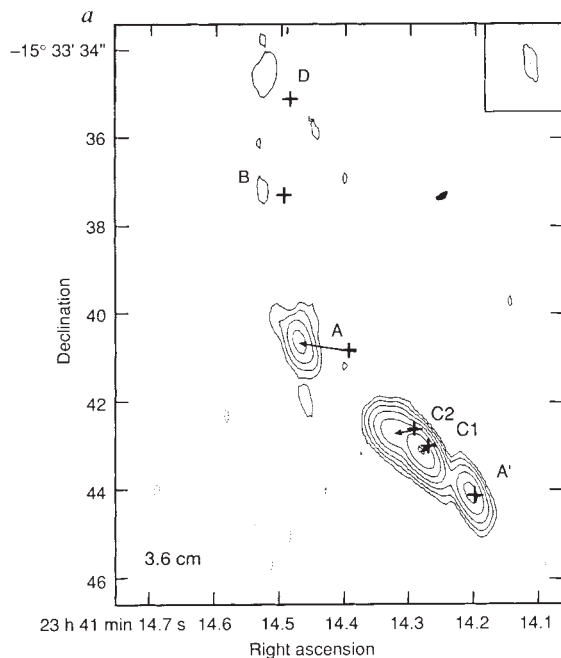


FIG. 1 The radio jet of R Aqr at $\lambda = 3.8$ cm and $\lambda = 6.2$ cm. The contours shown are −0.1, 0.1, 0.2, 0.4, 0.8, 1.6 and $3.2 \text{ mJy beam}^{-1}$. At the distance of 200 pc, 1 arcsec equals 200 au ($3 \times 10^{15} \text{ cm}$). The crosses show the positions of the components at previous epochs (see Table 1). The proper motion of components A', C1, C2 and A over the past ten years is shown by arrows in the 3-cm map, and the proper motion of components B and D over the past ten years are shown in the 6-cm map. The 1σ r.m.s. noise levels of the two maps are $31 \mu\text{Jy}$ and $33 \mu\text{Jy}$, respectively. A continuum bandwidth of 100 MHz was used. The beam sizes in the 3-cm and 6-cm maps are $0.66'' \times 0.31''$ (PA = 14°) and $1.11'' \times 0.53''$ (15°). The 50% contours of the dirty beam are shown in the upper right corners. We also made a 3-cm map using a 340-kilowavelength ($k\lambda$) taper (not shown). Tapering

effectively weights the data by $w(b) = C \exp(-(b/b_0)^2)$, where b is the baseline, b_0 is the taper and C is a scaling constant. The tapered map has a lower resolution than an untapered one. The taper in our 3-cm map was chosen so that the resolution would be very similar to the 6-cm map, enabling us to make a direct comparison between the two wavelengths. The dirty beam in the tapered 3-cm maps was $1.03'' \times 0.53''$ (10°). A CLEAN beam equal to the 6-cm dirty beam was used for this map. The structure seen in this map is essentially the same as in the 6-cm map. In addition, we made a 6-cm map tapered to $100 k\lambda$, but it revealed no new information. The size of all maps was 256×256 pixels, with a pixel size of $0.1''$ in the untapered 3-cm map and $0.15''$ in the other maps.

TABLE 2 Proper motions of the components in the jet of R Aqr

	RA (s yr ⁻¹)	Dec ('' yr ⁻¹)	Absolute ('' yr ⁻¹)	PA (°)	Relative ('' yr ⁻¹)	PA (°)	Inclination (°)
A'	+0.0031(22)	+0.005(48)	0.019(32)	83(61)	0.030(36)	28(71)	—
C1	+0.0021(5)	-0.021(11)	0.037(8)	125(15)	0.000	—	—
C2	+0.0052(5)	-0.005(13)	0.075(8)	93(10)	0.047(16)	69(20)	-24.3
A	+0.0080(9)	+0.024(17)	0.118(13)	78(8)	0.096(19)	62(11)	-31.7
B	+0.0119(12)	+0.011(20)	0.173(17)	86(7)	0.145(22)	77(9)	-21.8
D	+0.0102(27)	+0.095(45)	0.175(41)	57(14)	0.165(43)	45(15)	-21.1

The estimated 1σ accuracy of the fit is shown in parentheses. The first four columns list the absolute proper motion, and the position angle. The next two columns give the proper motion relative to the core, C1, and the relative position angle and the last column gives the inclination of the jet components to the plane of sky: 0° is in the plane of sky, and 90° is towards the observer. Note: $10 \text{ km s}^{-1} \approx 0.0105'' \text{ yr}^{-1} \approx 0.00073 \text{ s yr}^{-1}$ at 200 pc in R Aqr.

arcsec per year and d is the distance in arcsec from the core (C1). Optical line observations¹ imply that typical radial velocities for components A, B and D with respect to the central star are about -56 , -55 and -60 km s^{-1} respectively. We also adopt a radial velocity of -20 km s^{-1} for C2. Combined with the proper motion, this implies observed space velocities equal to 48.8, 106.9, 148.5 and 167.4 km s^{-1} , and an inclination of $\sim 24^\circ$ from the plane of sky for C2, A, C and D, respectively. The dependence of the space velocity on the distance now becomes $v \approx 40.9 \text{ km s}^{-1} \sqrt{d/100 \text{ AU}}$. This could be accounted for by a constant acceleration of $4.66 \times 10^{-3} \text{ cm s}^{-2}$ over a distance of $\sim 1,400 \text{ AU}$. From the observed distance to C2, we can now estimate that the components emerged from the core 29.4, 65.1, 95.6 and 106.8 years ago. The time difference between the emergence epochs of C2 and A in this model is 35.7 years. The difference between the emergence epoch of A and the average of B and D is 36.1 years, suggesting that in this model also B and D are really a single component which has broken in two. This timescale is shorter than the suggested optical period of 44 yr (ref. 23). This simple model has difficulty in explaining simultaneously the shape of the jet and the fact that the position angle of the proper motion is systematically $\sim 15^\circ$ larger than that of the radius vector to a component. The discrepancy is even larger when compared with the apparent local direction of the jet. In a continuous jet one would expect the proper motion vectors to align with the jet as one moves along the jet. This inconsistency cannot be eliminated by adding a constant initial velocity to the components.

In situ acceleration may also play an important part in diverting the momentary velocity vector from the preferred direction along the jet. For example, if the jet follows magnetic field lines near R Aqr, then a twisted helical field, with a period equal to the binary period, could explain the synchronous proper-motion offsets if the components emerge at intervals of one binary period. The possibility that the components are slowly moving shocks in a very fast streaming plasma will be considered elsewhere (D.R.H.J. and H.J.L., manuscript in preparation).

The components in jets from active galactic nuclei detected by very-long-baseline interferometry seem to follow each other's path: as a younger component reaches the position of an older one, it seems to follow the old path²⁵. This is also the case for the radio jet in SS433 (ref. 26), the only other galactic jet in which proper motion has been observed. We do not know if this also happens in the R Aqr jet. We clearly need a third epoch, after 1995, to establish whether the components travel on terminal 'bullet' velocities or whether their proper motion vectors will change direction to follow the path of the older component. In addition, on a timescale of $\sim 30 \text{ yr}$ we will have a chance of studying the breakup of the already elongated component A. $\square$

- Whitelock, P. A. *Publ. astr. Soc. Pacif.* **99**, 573-591 (1987).
- Gregory, P. C. & Seaquist, E. R. *Nature* **247**, 532-534 (1974).
- Bowers, P. F. & Kundu, M. R. *Astr. J.* **84**, 791-794 (1979).
- Ghigo, F. D. & Cohen, N. L. *Astrophys. J.* **245**, 988-991 (1981).
- Herbig, G. H. *IAU Circ. No.* 3535 (1980).
- Wallerstein, G. & Greenstein, J. L. *Publ. astr. Soc. Pacif.* **92**, 275-283 (1980).
- Kafatos, M., Hollis, J. M. & Michalitsianos, A. G. *Astrophys. J.* **267**, L103-107 (1983).
- Hollis, J. M., Kafatos, M. & Michalitsianos, A. G. & McAlister, H. A. *Astrophys. J.* **289**, 765-773 (1985).
- Hollis, J. M., Michalitsianos, A. G., Kafatos, M., Wright, M. C. H. & Welch, W. J. *Astrophys. J.* **309**, L53-57 (1986).
- Paresce, F., Burrows, C. & Horne, K. *Astrophys. J.* **329**, 318-325 (1988).
- Burgarella, D. & Paresce, F. *Astrophys. J.* **370**, 590-596 (1991).
- Lépine, J. R. D., LeSqueren, A. M. & Scalise, E. Jr *Astrophys. J.* **225**, 869-879 (1978).
- Zuckerman, B. *Astrophys. J.* **230**, 442-448 (1979).
- Spencer, J. H. et al. *Astr. J.* **86**, 392-409 (1981).
- Mauron, N., Nieto, J. L., Picat, J. P., Lelièvre, G. & Sol, H. *Astr. Astrophys.* **142**, L13-15 (1985).
- Hollis, J. M., Oliversen, R. J., Kafatos, M., Michalitsianos, A. G. & Wagner, R. M. *Astrophys. J.* **377**, 227-234 (1991).
- Hollis, J. M. et al. *Astrophys. J.* **361**, 663-666 (1990).
- Smithsonian Astrophysical Observatory *Star Catalogue* (US Government Printing Offices, Washington DC, 1966).
- Kafatos, M., Michalitsianos, A. G. & Hollis, J. M. *Astrophys. J. Suppl.* **62**, 853-874 (1986).
- Willson, L. A., Garnavich, P. & Mattei, J. A. *IAU Bull. Variable Stars*, No. 1961 (1981).
- Paresce, F. et al. *Astrophys. J.* **369**, L67-70 (1991).
- Zensus, A. in *Variability in Blazars* (Cambridge University Press, in the press).
- Morgan, B. *Ann. Rev. Astr. Astrophys.* **22**, 507-537 (1984).
- Kafatos, M., Hollis, J. M., Yusef-Zadeh, F., Michalitsianos, A. G. & Elitzur, M. *Astrophys. J.* **346**, 991-996 (1989).

ACKNOWLEDGEMENTS. We thank the NRAO AOC staff for their hospitality and help. NRAO is operated by Associated Universities, Inc., under agreement with the NSF.

A search during the 1991 solar eclipse for the infrared signature of circumsolar dust

Klaus-Werner Hodapp*, R. M. MacQueen†
& Donald N. B. Hall*

* Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

† Department of Physics, Rhodes College, 2000 North Parkway, Memphis, Tennessee 38112, USA

THEORETICAL suggestions¹⁻³ that there should be a ring of dust in near-ecliptic orbit about the Sun were supported by observations, during a total solar eclipse on 12 November 1966, of enhanced infrared emission⁴⁻⁶ from the solar corona at a distance of $4 R_\odot$ from the Sun's centre. The infrared emission was attributed to the sublimation of dust grains as they spiral into the Sun because of the Poynting-Robertson effect, by which solar radiation creates a tangential drag. Two months after the 1966 eclipse, the feature at $4 R_\odot$ was seen again in observations from a stratospheric balloon-borne coronagraph, as were additional features at 3.5, 8.7 and $9.2 R_\odot$ (ref. 6). Observations since then, however, have failed unambiguously to corroborate the earlier observations. We searched for excess infrared emission in the solar equatorial plane during the 11 July 1991 eclipse, using a wide-angle infrared camera on Mauna Kea, but failed to find any signature of dust evaporation. We argue that the earlier observations were credible, and therefore

Received 30 September; accepted 21 November 1991.

- Solf, J. & Ulrich, H. *Astr. Astrophys.* **148**, 274-288 (1985).
- Sopka, R. J., Herbig, G., Kafatos, M. & Michalitsianos, A. G. *Astrophys. J.* **258**, L35-39 (1982).
- Spergel, D. N., Giuliani, J. L. Jr & Knapp, G. R. *Astrophys. J.* **275**, 330-341 (1983).

that the circumsolar dust ring is a transient feature, perhaps due to the injection of dust into near-solar space by a Sun-grazing comet.

In 1970, a ground-based eclipse observation of the 9–10- μm -wavelength emission of the dust feature at $4 R_{\odot}$ was thwarted by variations in the thermally emitted sky background⁷. From multi-wavelength observations (I, H and K) at the 1970 eclipse, however, Peterson⁸ deduced an emission temperature of $\sim 2,100$ K for particles at $4 R_{\odot}$. During the 1973 solar eclipse, extensive airborne measurements in the 9–10- μm spectral region found evidence for enhanced emission near $4 R_{\odot}$, and the spectrum obtained was tentatively attributed to that arising from a silicate-type material⁹. Confusingly, however, this 10- μm radiance⁹ is roughly a factor of 5 greater than the upper limit determined from the 1970 observations⁷.

Later out-of-eclipse observations¹⁰ (1978) at wavelengths of 2.2 and 3.5 μm failed to detect any indication of thermal emission from coronal dust; additional unsuccessful attempts during eclipses in 1978 and 1980 have been reported^{11,12}. A complex, balloon-borne experiment^{13–15} at the eclipse of 1983 suggested that enhanced near-infrared emission (wavelengths 1.25, 1.65, 2.25 and 2.8 μm) and a polarization excess at 0.8 μm wavelength were present between 4 and 5 $R_{\odot}$, but the observed infrared emission was asymmetric and more extended than that observed in 1966 and 1967.

The original observations of emission peaks in the infrared coronal signal stimulated theoretical work on the nature and dynamics of near-solar interplanetary dust, work which has gone far beyond the initial considerations^{2,3} of straightforward absorption of sunlight and subsequent infrared emission by heated dust. Kaiser¹⁶ found that the four emission peaks discovered in 1966 and 1967 could be matched in spatial position by considering two types of refractory materials, with different complex indices of refraction: 'clean' and 'dirty' quartz. Mukai and collaborators^{17,18} developed the notion¹⁹ that as dust sublimation commences and the grain size decreases, the grains experience enhanced radiation pressure, halting their spiralling towards the Sun. They found^{20,21} that particles could become stabilized against Poynting–Robertson drag over 100–1,000 orbital revolutions of a dust grain and concentrate in region(s) near the sublimation zone(s). Lamy²² confirmed this with a

numerical simulation. The general nature of emission peaks could therefore be explained by this concentration of dust. On the other hand, refs 17 and 18 have found that the actual residence time for a sublimating grain depends upon the temperature of the grain and that some materials (silicate is a good example) show a strong temperature dependence with radius. Thus, different particle sizes sublime at different distances from the Sun, implying that it is more difficult to achieve a concentration of emitting dust in a restricted spatial region. As a result, it is suggested that silicate is unlikely to be responsible for the presence of any coronal emission peak due to sublimation¹⁸.

After 25 years of studies, therefore, we are left with an unclear picture of the nature, or perhaps even the existence, of a signature of thermal emission from circumsolar dust. In this uncertain context, new infrared observations were planned for the eclipse of 1991, exploiting the extended duration of totality, and the fact that the path of totality crossed the site of the Mauna Kea observatories, well known as a superior site for infrared astronomical observations.

The experiment was designed to obtain infrared images in the 1–2.5- μm region between 2.5 and 15 $R_{\odot}$ while minimizing scattering in the instrument optics. We used a 256×256 HgCdTe detector array developed by the Rockwell International Science center for NASA's NICMOS (near infrared camera and multi-object spectrometer) project, in the University of Hawaii's infrared camera (K.W.H. *et al.*, submitted to *Publ. Astr. Soc. Pacific*). A CaF_2 objective lens replacing the telescope optics normally used, gave this camera an image scale of 1.66 arcmin per pixel and $\sim 7^\circ \times 7^\circ$ field of view. To reduce stray light and potentially confusing multiple reflection images from the bright inner corona, an external serrated occulting disk of 49 mm diameter was placed 2.5 m in front of the objective, covering the inner 2.5 $R_{\odot}$. The whole instrument was mounted in the tube of one of the University of Hawaii 0.6-m telescopes on Mauna Kea, with the primary mirror of this telescope covered and the secondary removed.

During totality, we obtained images in narrow-band filters centred at 2.12 μm and 2.36 μm as well as in broad-band filters centred at 1.2, 1.65 and 2.2 μm . Polarization data were obtained at 1.6 μm and 2.2 μm , through a set of three wire-grid polarizers. The narrow-band filter observations were intended to assist in discriminating between possible infrared emission from the cool ($T < 230$ K) stratospheric volcanic dust from the June 1991 eruption of Mt Pinatubo on the Philippine Island and anticipated thermal emission from the hotter ($T > 1,000$ K) circumsolar dust. The weather conditions during totality were poor. Volcanic dust scattering and a near-solid layer of clouds below the summit led to very substantial scattering in the atmosphere and thus a bright sky background. But this scattering was spatially uniform, and no infrared emission from volcanic dust could be identified. More of a problem was caused by isolated thin cirrus clouds that drifted through the field of view during totality. Unfortunately, many of our 15 frames obtained during totality are degraded by this nonuniform sky brightness. Our clearest image was the last of our sequence, a 10-s exposure through the 2.12- μm narrow-band filter ending 4 s before the end of totality (Fig. 1). Although scattering off the cirrus cloud affects the analysis of faint signals near the sky level, infrared extinction by the cirrus was less of a problem, so that the bright inner parts of the corona can be studied adequately. Because of a mechanical problem in our finderscope, our external occulting disk was not perfectly aligned with the centre of the Sun, resulting in the contours of the corona being offset from the disk centre (Fig. 2a). This offset was small enough that the disk still covered the brightest parts of the inner corona and that the stray light was at the expected levels.

Figure 1 clearly shows electron-scattered coronal streamers extending to $\sim 10 R_{\odot}$ from the Sun centre. The outer corona appears flattened in (or near) the ecliptic plane. Some scattering from cirrus clouds is visible in the northwestern corner of the

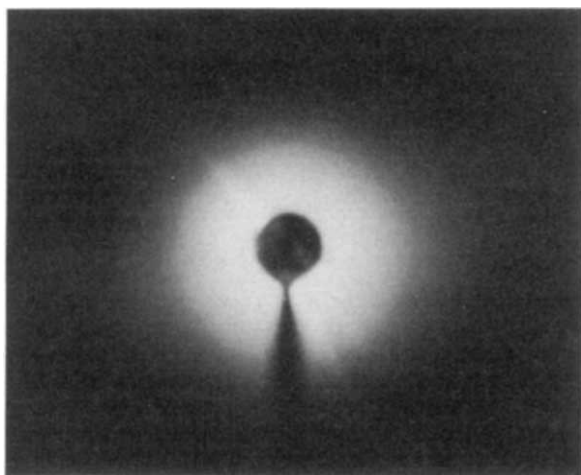


FIG. 1 Logarithmic grey-scale image of the outer corona at a wavelength of 2.12 μm . The field of view is $\sim 7^\circ \times 7^\circ$. North is up and west is right in this image. The data frames taken during totality were flatfield-corrected using frames of the daytime sky taken under cloud-free conditions. Sky-frames taken in the evening after sunset and having a signal level comparable to the conditions during the solar eclipse were subtracted from the frames. Known bad pixels of the NICMOS3 detector array ($\sim 2\%$ of all pixels) were replaced by a two-dimensional linear interpolation through the neighbouring pixels. The out-of-focus image of the external occulting disk and its support arm are visible in the lower part of the image.

image. Otherwise, no spatial structure is visible; in particular, no ring segments or other spatial structure at the radial position(s) of the previously reported dust emission features around the Sun. Figure 2a is a contour plot of the same data, clearly showing the flattening of the outer corona in (or near) the ecliptic plane. Radial cross-sections through the inner corona (Fig. 2b) again appear completely featureless. If any thermal emission features were present in the corona at the time of the 1991 eclipse, they must have been much less conspicuous than those previously reported. Specifically, our upper limit for the radiance of any potential signature of thermal emission is about a factor of 20 less than the radiance of the $4 R_{\odot}$ feature reported⁶ at the 1966 eclipse.

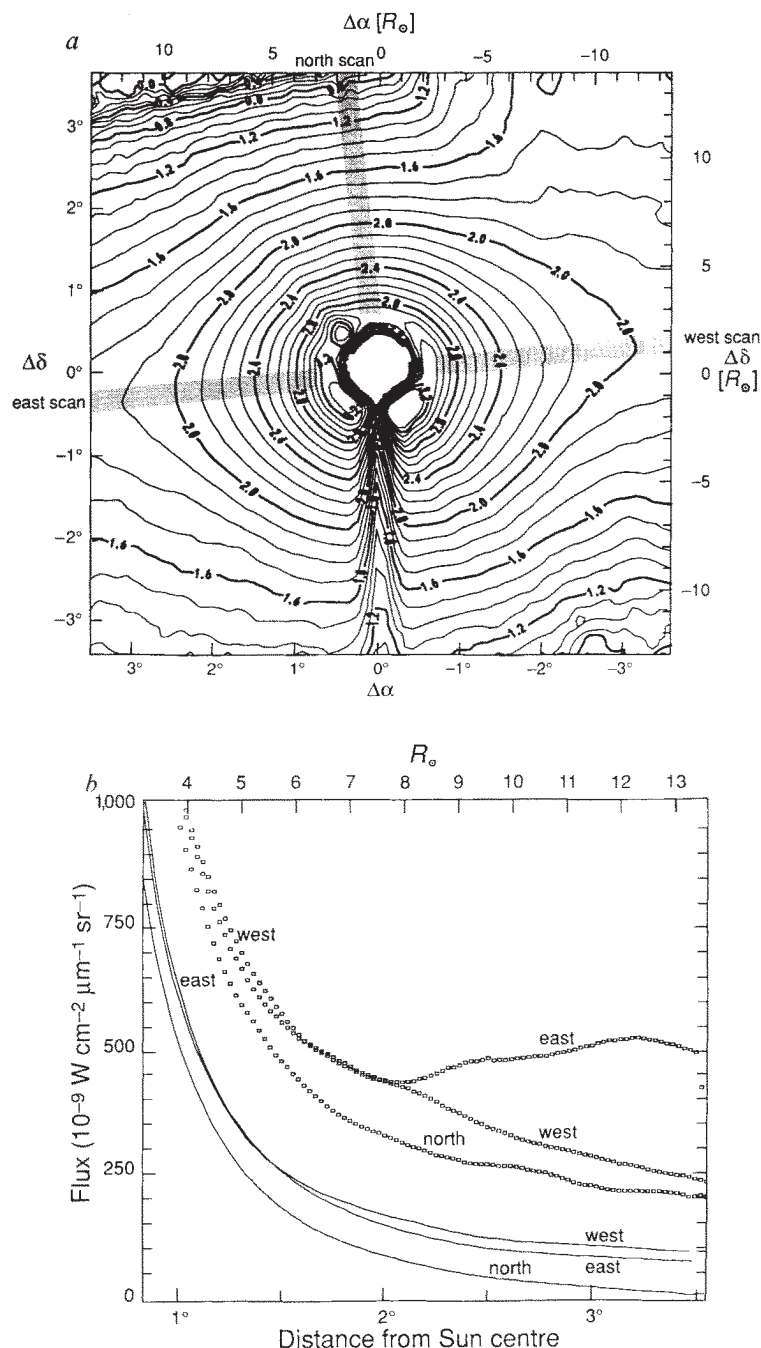
This lack of detection of any spatial feature in the infrared corona suggests either that previous identification of features in the corona in the infrared spectral region was incorrect, with the structures so identified resulting, perhaps, from instrumental artefacts; or that the presence of circumsolar signatures of dust

vaporization is temporally variable, perhaps associated with heretofore unidentified variations in the near-solar dust magnitude or composition.

We have no reason to doubt the reality of the previous observations; the reported signals of the inner coronal feature attributed to thermal emission at the 1966 eclipse^{4,6} and shortly thereafter⁶ were unambiguous. As there have been several positive identifications of enhanced dust emission near the Sun, using different observing systems, it seems unlikely that instrumental artefacts could account for the reported signatures.

Therefore, we must turn to the possibility that circumsolar dust is temporally variable in magnitude or composition. Neither the timescale for short-period cometary capture nor that for asteroidal collision influx is consistent with variations of the near-solar dust population. The timescale for the capture of short-period comets by Jupiter is ~ 500 – $1,000$ years, assuming estimates for the frequency of new comet appearances and capture probability²³. The timescale for asteroidal collision of

FIG. 2 a, Contour plot of the $2.12\text{-}\mu\text{m}$ image of the outer corona (Fig. 1) in logarithmic scale in units of $\log_{10} B_{\lambda}/(10^{-9} \text{ W cm}^{-2} \mu\text{m}^{-1} \text{ sr}^{-1})$ (where B_{λ} is the surface brightness, and α and δ are the right ascension and declination on the sky). Because of the high sky brightness, we could not obtain photometric calibration of our system during the eclipse or immediately after it. Our calibration relies on night-time observations of the standard stars α Boo and α Lyr, taken under photometric conditions on 16 July 1991 with our eclipse instrument, and on absolute photometry of these stars³⁸. These measurements were still affected by the volcanic dust, which was nearly uniformly distributed. Broad-band K measurements taken in a photometric night after the eclipse (14 July 1991) at the IRTF Infrared Telescope Facility (K. Sellgren, personal communication) were used to give an estimate of K-band extinction of 0.08 mag per unit air mass, a value only slightly worse than usually measured under dust-free conditions. Although both the photometric calibration ($\pm 2\%$) and the measurement of the extinction are of higher precision, we estimate this value to have errors of the order $\pm 10\%$. This error comes mainly from the uncertainties about the stability of the atmospheric conditions between the eclipse and the calibration nights. The contours in the inner $50'$ are affected by the vignetting zone of the (out-of-focus) disk. The brightness distribution is flattened in the corona in or near the plane of the ecliptic. Note the offset between the disk centre and the corona contours caused by slight telescope mispointing during the eclipse. b, Radial line plots. The positions of these plots are indicated in a as shaded areas. Horizontal axis: degrees from the centre of the highest corona isophotes, identified as the centre of the eclipsed solar disk. Solid lines: cuts through the $2.12\text{-}\mu\text{m}$ image in Figs 1 and 2a, the last image obtained during totality, under the best sky conditions we had during totality. Open squares: identically located cuts through the first $2.12\text{-}\mu\text{m}$ image obtained during totality. At the beginning of totality, a thick band of cirrus stretched through the field of view, increasing the sky background due to scattering. This first image is a worst-case example of the effect of clouds on the detection of circumsolar features. Most importantly, the typical scale of cirrus clouds, including their drift during the integration time, was $\sim 30'$, much larger than the expected and previously observed extent of the circumsolar features. Single noisy pixels produce spikes of at most $5 \times 10^{-9} \text{ W cm}^{-2} \mu\text{m}^{-1} \text{ sr}^{-1}$. We conclude that on the previously observed⁶ scale of ~ 0.2 – $0.3 R_{\odot}$ ($3'$ – $5'$ or 2–3 pixels) for the circumsolar features, we would certainly have detected excess flux of $10 \times 10^{-9} \text{ W cm}^{-2} \mu\text{m}^{-1} \text{ sr}^{-1}$.



small (metre-sized) bodies is probably even longer²⁴. And finally, any such variations in dust production in the Solar System would be further filtered by the timescale for the infall of dust due to Poynting–Robertson drag: roughly 1 million years²⁵.

Excepting the presence of Sun-grazing comets^{26–28}, and possibly, large solar energetic events²⁹ (X-ray flaring), there is little evidence for temporal variations in the light scattered by the dust particles, which constitutes the F-corona and/or zodiacal light^{30,31}. The ‘direct’ cometary infusion of material into the near-solar regions by Sun-grazing comets seems to be a short-lived, transitory process, lasting only several tens of hours²⁶. It is now understood that Sun-grazing comets belonging to the so-called Kreutz group are highly episodic in their appearance²⁸. Further, Kreutz-group cometary dust input is spatially inconsistent with the concentration of the dust features in or near the ecliptic plane, although there are substantial uncertainties in, and differences between, the reported spatial distributions^{4,6,13}.

Still it seems that we must seek a near-solar influx of cometary dust to explain the past observations, especially those of 1966 and 1967 in which a clear signature of the dust was present. Thus, although visible light observations at the November 1966 eclipse revealed no anomalous enhancement of the F-coronal component^{32–34}, it is possible that an infusion of dust near the Sun would be undetected in the visible but noticeable in the infrared, as forward-scattering is somewhat insensitive to near-solar dust. Mass estimates for the dust detected by the earlier infrared experiments are uncertain by several orders of magnitude, because of the uncertainties in the input parameters³⁵. Still, we note that the mass of the dust detected earlier is consistent with cometary mass-loss rates near the Sun.

Satellite-borne coronagraphs, which have been effective in the detection of small members of the Kreutz group, were not available in 1966 and 1967. The possibility that a Kreutz-group member provided an additional source of dust near the time of the November 1966 eclipse or before cannot be ruled out. The prolific dust-emitting Kreutz group comet 1965 VIII (Ikeya-Seki) would be an excellent candidate. Even so, it is difficult to understand how such dust could have remained near the Sun for the many months necessary to explain the enhanced dust emission in the period from November 1966 to early January 1967. Lorentz forces on charged dust will be much more important in the dynamics of dust near the Sun than farther away, where the magnetic and electric fields are smaller. Such forces will vary in time, as we know that the large-scale solar magnetic field, and presumably the electric field, are significantly modified over timescales of hours to months^{36,37}. As yet, we are unaware of any calculations of dust dynamics in the complex and variable near-solar environment. □

Received 5 August; accepted 2 December 1991.

- Russell, H. N. *Astrophys. J.* **61**, 49–71 (1929).
- Over, J. *Proc. Kon. Neder. Akad. Wetensch.* **B51**, 74–84 (1958).
- Peterson, A. W. *Astrophys. J.* **138**, 1218–1230 (1963).
- Peterson, A. W. *Astrophys. J.* **148**, 137–39 (1967).
- Peterson, A. W. *Astrophys. J.* **155**, 1009–1015 (1969).
- MacQueen, R. M. *Astrophys. J.* **154**, 1059–1076 (1968).
- Markin, W. G., MacQueen, R. & Lee, R. H. *Astr. Astrophys.* **31**, 17–21 (1974).
- Peterson, A. W. *Bull. Am. Astr. Soc.* **3**, 500 (1971).
- Lena, P., Viala, Y., Hall, D. & Soufflot, A. *Astr. Astrophys.* **37**, 81–86 (1974).
- Mampaso, A. & Sanchez Magro, C. *Rev. Mexic. Astr. Astrofis.* **8**, 3–6 (1983).
- Zirker, J. *Total Eclipses of the Sun*, 121–122 (Van Nostrand Reinhold, city, 1984).
- Keller, C. F. & Liebenberg, D. *Los Alamos Sci.* **2**, 4 (1981).
- Mizutani, M., Maihara, T., Hiromoto, N. & Tamaki, H. *Nature* **312**, 134–136 (1984).
- Isobe, S., Hirayama, T., Baba, N. & Miura, N. *Nature* **318**, 644–646 (1985).
- Isobe, S., Hirayama, T., Baba, N. & Miura, N. *Publ. astr. Soc. Japan* **39**, 667–677 (1987).
- Kaiser, C. B. *Astrophys. J.* **159**, 77–92 (1970).
- Mukai, T. *Properties and Interactions of Interplanetary Dust* (eds Giese, R. & Lamy, Ph.) 9–62 (Reidel, Dordrecht, 1985).
- Mukai, T. & Schwehm, G. *Astr. Astrophys.* **95**, 373–382 (1981).
- Belton, M. J. *Science* **151**, 35–44 (1966).
- Mukai, T., Yamamoto, T., Hasegawa, H., Fujiwara, A. & Koike, C. *Publ. astr. Soc. Japan* **26**, 445–458 (1974).
- Mukai, T. & Yamamoto, T. *Publ. astr. Soc. Japan* **31**, 585–595 (1979).
- Lamy, Ph. L. *Interplanetary Dust and Zodiacal Light* (eds Elsässer, H. & Fechtig, H.) 437–442 (Springer, Berlin, 1976).
- Fernandez, J. A. & Ip, W.-H. *Comets in the Post-Halley Era* (eds Newburn, R. L. Jr., Neugebauer, M. & Rahe, J.) 487–535 (Kluwer, Dordrecht, 1991).

- Dohnanyi, J. S. *Cosmic Dust* (ed. McDonnell, J. A. M.) 527–605 (Wiley, Chichester, 1978).
- Robertson, H. P. *Mon. Not. R. astr. Soc.* **97**, 423–438 (1937).
- Michels, D. J., Sheeley, N., Howard, R. & Koomen, M. J. *Science* **215**, 1097–1102 (1982).
- Sheeley, N., Howard, R., Koomen, M. J. & Michels, D. J. *Nature* **300**, 239–242 (1982).
- MacQueen, R. M. & St Cyr, O. C. *Icarus* **90**, 96–106 (1991).
- Blackwell, D. E. & Ingham, M. F. *Mon. Not. R. astr. Soc.* **122**, 143–155 (1961).
- Misconi, N. Y. *Astr. Astrophys.* **51**, 357–365 (1976).
- Leinert, C. & Pitz, E. *Astr. Astrophys.* **210**, 399–402 (1989).
- Newkirk, G. A., Dupree, R. G. & Schmahl, E. J. *Solar Phys.* **15**, 15–39 (1970).
- Waldmeier, M. Z. *Astrophys.* **69**, 193–213 (1968).
- Arnqvist, W. & Menzel, D. H. *Solar Phys.* **11**, 82–91 (1970).
- Giese, R. H. & Grün, E. *Interplanetary Dust and Zodiacal Light* (eds Elsässer, H. & Fechtig, H.) 135–139 (Springer, Berlin, 1976).
- MacQueen, R. M. *The Sun and the Heliosphere in Three Dimensions* (ed. Marden, R. G.) 5–18 (Reidel, Dordrecht, 1985).
- Low, B. C. *Ann. Rev. Astr. Astrophys.* **28**, 491–524 (1990).
- Strecker, D. W., Erickson, E. F. & Witteborn, F. C. *Astrophys. J. Suppl. Ser.* **41**, 501–512 (1979).

ACKNOWLEDGEMENTS. We thank T. Brown, G. Ching, J. Rayner and A. Whinery for contributions to this work, and B. Greeley and K. LaCount for assistance in obtaining the observations. R.M. acknowledges support from the NSF.

Cellular convection in the atmosphere of Venus

R. David Baker II* & Gerald Schubert*†

* Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA

† Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California 90024, USA

AMONG the most intriguing features of the atmosphere of Venus is the presence of cellular structures near and downwind of the subpolar point. Ultraviolet images from the Mariner 10 and Pioneer Venus missions at cloud-top heights of 60–70 km indicate that these cells are chiefly found only during the afternoon at low latitudes, with dark-rimmed cells having horizontal scales of roughly 500–1,000 km, and bright-rimmed cells having horizontal scales of roughly 200–500 km (refs 1, 2). Images of the subsolar region from the recent Galileo flyby of Venus confirm these earlier observations—cellular features with horizontal scales of 250–1,000 km have been observed³ (Fig. 1a, b). It has been suggested that the structures are atmospheric convection cells^{2–5}, and, in fact, numerical simulations of three-dimensional compressible convection produce features reminiscent of the cellular structures found in the subsolar Venus atmosphere (Fig. 1c)⁶. There has been some difficulty, however, in accounting for the sizes of these structures. The convection cells were thought to be located in a neutrally stable cloud layer, roughly 50 km above the surface, with a thickness of only a few kilometres. So broad and thin a convection cell would pose a severe challenge to the dynamics of convection. Here we propose that strongly penetrative convection into the stable regions above and below the neutrally stable cloud layer coupled with penetrative convection from the surface increases the vertical dimensions of the cells, thereby helping to explain their large horizontal extent.

Vertical profiles of static stability for the Venus atmosphere have been determined from pressure and temperature data obtained by the Pioneer Venus probes and the Vega 2 lander^{7,8}. Figure 2 presents static stability profiles at low latitudes from the Pioneer Venus and Vega 2 missions. The profiles for the probes are strikingly similar, with neutrally stable layers found ~20–30 km and 50–55 km above the surface and a stable layer between them. Neither the Pioneer Venus probes nor the Vega 2 lander sampled the subsolar region. Veneras 9, 10, 11 and 12 did, however, measure temperature and pressure profiles near the subsolar point⁹. Stability profiles derived from data from Veneras 10, 11 and 12 show the basic stability structure found by Pioneer Venus and Vega 2, with the notable exceptions that the lower adiabatic layer extends to the surface, and Venera 10 measured a slightly unstable region from 30 to 35 km. Radio occultation measurements with Veneras 9 and 10 indicate that the stable layer is sometimes small or even absent, and the upper

neutrally stable layer reaches as high as 60 km above the surface¹⁰.

Thus, in the subsolar region, an adiabatic layer probably extends from the surface to ~ 30 km, in agreement with one-dimensional radiative-convective models¹¹. A second adiabatic layer exists ~ 50 – 60 km above the surface. These layers are probable sites of thermal convection. A stable layer, somewhat weakened in the subsolar region, separates the two convective layers¹⁰. Here we argue that convection in the two adiabatic layers may strongly penetrate into the weakened intermediate stable layer and possibly form whole-atmosphere convection. This whole-atmosphere convection could help to explain the existence of the cellular structures found in the subsolar region.

Convection in the cloud-level adiabatic layer (about 10 km thick) has been suggested to be responsible for the cellular features found in the subsolar region¹². But with cellular length scales as large as 1,000 km, convection in this layer implies aspect ratios (horizontal dimension/vertical dimension) as high as 10^2 . By comparison, laboratory experiments involving an internally heated layer have produced horizontal scales only five times larger than the vertical scale¹³ and mesoscale cellular convection in the Earth's atmosphere generates aspect ratios of about ten¹⁹. Anisotropic eddy diffusion has been proposed to explain the large aspect ratios of cloud-level convection cells on Venus⁵, but the study involved a linearized, Boussinesq fluid and did not account for a fully nonlinear, compressible regime. Whole-atmosphere convection in the subsolar region of Venus offers a more likely reconciliation of this aspect ratio problem because, as discussed below, the aspect ratio of convection cells that extend from the surface of Venus to the cloud tops is comparable to the aspect ratios of terrestrial mesoscale convection cells and convection cells in the laboratory.

Studies of two-dimensional compressible convection have shown that subtle yet significant differences exist between compressible and Boussinesq convection. In an anelastic model (a model that filters out sound waves while retaining quasi-static effects of compressibility), downward convective penetration into an adjacent stable layer is substantially greater than for a Boussinesq system^{14,15}. Similar results were achieved for a fully nonlinear, compressible system, with the downward penetration into the stable layer as deep as 0.58 pressure scale heights¹⁶. For a two-layer system with an unstable layer bounded above by a stable layer, upward penetration into the stable layer can be as large as 1.28 pressure scale heights¹⁶. Thus we see that nonlinear, compressible convection penetrates deeply into surrounding stable layers. Using a scale height of 10 km for 30 km above the surface of Venus¹⁷, we find that upward penetration from the lower convective layer could extend as much as 12.8 km into the stable layer. Similarly, with a scale height of 7.5 km at 50 km above the surface¹⁷, convection from the upper adiabatic layer could penetrate as deep as 4.4 km into the stable layer. We believe this convective penetration may contribute significantly to the cellular structure found in the subsolar region. Furthermore, two-dimensional simulations of fully compressible convection spanning multiple scale heights indicate that convective motions tend to organize in cells that extend over the full height of the layer rather than in successive rolls in the vertical as assumed in mixing-length theory¹⁸. Coupled with the weakened stable layer in the subsolar region and the likelihood of deep penetrative convection, this multiple scale height structure may help explain the large horizontal scales of the cellular features found in the subsolar region. In effect, convection may be occurring on a vertical scale from the surface of Venus to the cloud tops, implying aspect ratios of ~ 10 for convection

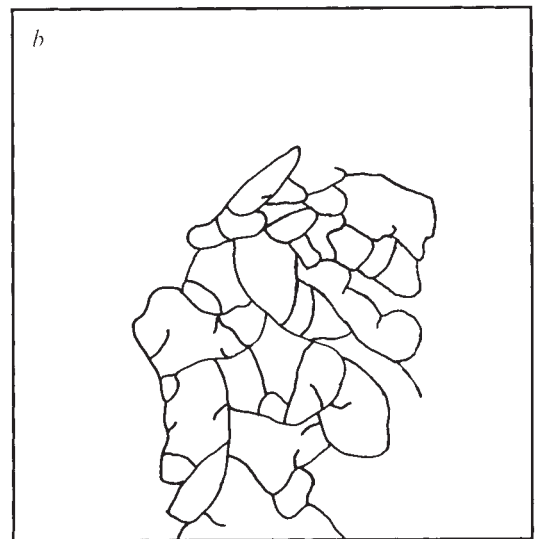
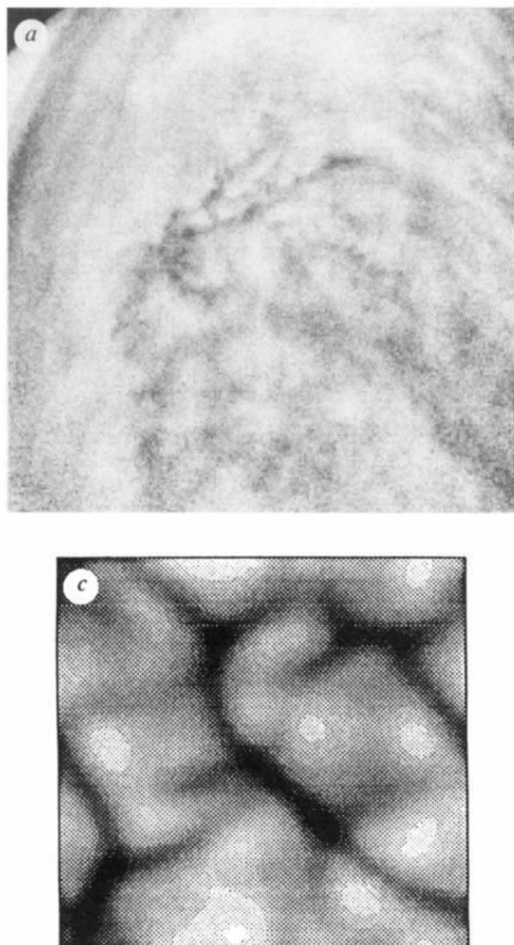
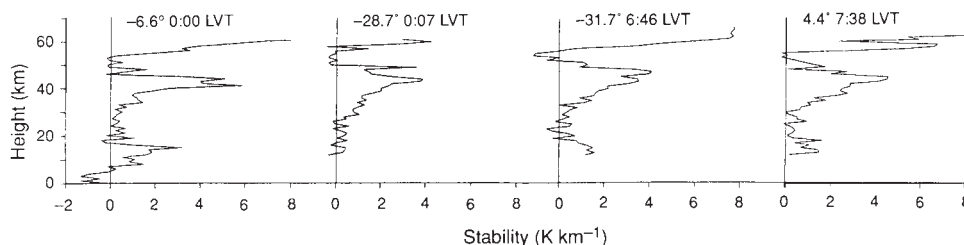


FIG. 1 *a*, Ultraviolet image of cellular convection in the subsolar region of Venus (Galileo image 1892460, effective wavelength 418 nm). The image size is $\sim 90^\circ$ of latitude by 80° of longitude. The equator runs vertically across the image, with the image centre located at 4° latitude and 37° west (downwind) of the subsolar point. North points toward the left edge of the image. The subsolar point is located at the upper edge, with the retrograde motion of the atmosphere moving from top to bottom. The local Venus time of the image centre is 14:28. Dark-rimmed cellular features have horizontal scales of 250–1,000 km. *b*, Sketch of the Galileo image tracing the boundaries of the cellular structures found in the subsolar region. *c*, Results from a three-dimensional numerical simulation of compressible convection with Prandtl number of unity. Upward motions are displayed as bright regions, downward motions as darker regions⁶. The structures seen here closely resemble the cellular features found in the subsolar region of Venus.

FIG. 2 Static stability plotted against height in the low-latitude atmosphere of Venus from the Vega 2 lander (far left), Pioneer Venus (PV) night probe, PV day probe, and PV sounder. The latitude and the local Venus time (LVT) of entry are given for each probe.



cells. These ratios closely agree with the aspect ratios of meso-scale cellular atmospheric convection on Earth¹⁹. Further, whole-atmosphere convection would tend to stabilize convection planforms against distortion by the vertical shear of the retrograde zonal circulation, because the base of a convection cell would be in the deep, slowly circulating part of the atmosphere where vertical shear is relatively small. Of course, the possibility remains that convection from the two neutrally stable layers does not penetrate strongly enough to form whole-atmosphere convection; still, the penetration should be large enough to influence the dynamics in the stable layer and quite possibly the structure of the cellular features.

The formation of whole-atmosphere convection may introduce significant dynamical effects in the subsolar region. As convection penetrates into the intermediate stable layer, internal gravity waves generated by the penetration may transport energy and momentum to the adjacent convective layers. This communication between the two convective layers by internal gravity waves could be the vital link in the formation of whole-atmosphere convection. In addition, whole-atmosphere convective motions may force vertically trapped, horizontally propagating waves in the stable layer outside the subsolar region; it has been shown that the stable layer is capable of supporting such waves²⁰. Convective activity has also been suggested as the likely candidate for forcing planetary-scale waves¹⁷, and both penetrative and whole-atmosphere convection may be principal wave-generating mechanisms in the subsolar region. Thus the presence

of deeply penetrative convection and whole atmosphere convection may strongly influence the general circulation of the Venus atmosphere, making future investigations into the formation of whole-atmosphere convection a worthwhile endeavour. □

Received 18 September; accepted 2 December 1991.

1. Murray, B. C. *et al.* *Science* **183**, 1307–1315 (1974).
2. Rossow, W. B. *et al.* *J. geophys. Res.* **85**, 8107–8128 (1980).
3. Belton, M. J. S. *et al.* *Science* **253**, 1531–1536 (1991).
4. Belton, M. J. *et al.* *J. atmos. Sci.* **33**, 1394–1417 (1976).
5. Covey, C. C. & G. Schubert, *Nature* **290**, 17–20 (1981).
6. Toomre, J. *et al.* *Comp. Phys. Comm.* **59**, 105–117 (1990).
7. Seiff, A. *et al.* *J. geophys. Res.* **85**, 7903–7933 (1980).
8. Linkin, V. M. *et al.* *Cosmic Res.* **25**, 501–512 (1987).
9. Avduvskiy, V. S. *et al.* *Venus* (eds Hunten, D. M., Colin, L., Donahue, T. M. & Moroz, V. I.) 280–298 (University of Arizona Press, 1983).
10. Seiff, A. *Venus* (eds Hunten, D. M., Colin, L., Donahue, T. M. & Moroz, V. I.) 215–279 (University of Arizona Press, 1983).
11. Pollack, J. B., Toon, O. B. & Boese, R. *J. geophys. Res.* **85**, 8223–8231 (1980).
12. Schubert, G. *et al.* *J. geophys. Res.* **85**, 8007–8025 (1980).
13. Tritton, D. J. *Nature* **257**, 110–112 (1975).
14. Massaguer, J. M. *et al.* *Astr. Astrophys.* **140**, 1–16 (1984).
15. Zahn, J.-P., Toomre, J. & Latour, J. *Geophys. Astrophys. Fluid Dynam.* **22**, 159–193 (1982).
16. Hurlbert, N. E., Toomre, J. & Massaguer, J. M. *Astrophys. J.* **311**, 563–577 (1986).
17. Schubert, G. *Venus* (eds Hunten, D. M., Colin, L., Donahue, T. M. & Moroz, V. I.) 681–765 (University of Arizona Press, 1983).
18. Hurlbert, N. E., Toomre, J. & Massaguer, J. M. *Astrophys. J.* **282**, 557–573 (1984).
19. Agee, E. M., Chen, T. S. & Dowell, K. E. *Bull. Am. met. Soc.* **54**, 1004–1012 (1973).
20. Schinder, P. J. *et al.* *J. atmos. Sci.* **47**, 2037–2052 (1990).

ACKNOWLEDGEMENTS. We thank M. J. S. Belton and the Solid State Imaging Team for furnishing the Galileo image, and J. Toomre for the figure of the 3-D numerical simulation of compressible convection. This research was supported by NASA.

Intercalation of solid C₆₀ with iodine

Qing Zhu^{*†}, David E. Cox[‡], John E. Fischer^{*†}, Krzysztof Kniaz^{*†}, Andrew R. McGhie^{*} & Otto Zhou^{*†}

^{*} Laboratory for Research on the Structure of Matter, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6272, USA

[‡] Department of Physics, Brookhaven National Laboratory, Upton, New York 11973, USA

METALLIC and superconducting donor-type intercalation compounds of C₆₀ are now well established. These involve electron transfer from a sublattice of alkali-metal dopants to a sublattice of fullerene molecules. Three stoichiometric phases have been identified^{1–3}, and a binary phase diagram describing composition regions of two-phase coexistence has been proposed². Oxidative intercalation by electron acceptors, meanwhile, has not previously been reported. Ohno *et al.*⁴ recently presented photoemission data suggesting that solid C₆₀ will take up only minor amounts of iodine, resulting in a non-metallic product which is not a stoichiometric compound⁴. Here we report that, by using different reaction conditions, we have prepared a highly crystalline C₆₀:

iodine compound of ideal stoichiometry C₆₀I₄, with an alternating guest–host layer structure closely resembling that of classic intercalation compounds. The 300-K resistivity exceeds 10⁹ Ω cm and we observe no superconductivity down to 4 K, despite the fact that the inter-fullerene separation lies in the range of that in the superconducting M₃C₆₀ phases (where M is K, Rb, Cs). C₆₀I₄ is apparently the first example of a fullerite intercalation compound in which there is no electron transfer between C₆₀ and the intercalate.

Discussion of the possibility of a superconducting phase in iodine-doped C₆₀ (ref. 5) invoked the idea of acceptor-type fullerene ‘salts’, resulting for example from oxidative intercalation by halogens. This would seem to be ruled out *a priori* by the large C₆₀ molecular ionization potential⁶ and the absence of reversible oxidation steps in solution electrochemistry⁷, although the electrostatic penalty per molecule could be partly offset by the Madelung energy of a resulting ionic lattice. On the other hand, there are many examples of synthetic metals obtained by halogen doping into other carbon-based hosts. Iodine is the dopant favoured for testing the effect of polymer morphology on the conductivity of polyacetylene compounds⁸; bromine, and the inter-halogen IBr and ICl intercalate readily into graphite at 300 K either from the vapour or when graphite is immersed in halogen-CCl₄ solutions⁹. These are generally p-type metals, and counter-ions such as I₃[−] and I₅[−] have been identified¹⁰.

Vapour-phase reactions of iodine with pure C₆₀ were carried out at 250 °C for several days in evacuated pyrex tubes, using starting mixtures with ratios I:C₆₀ ≈ 1, 2, 3 and 12. The C₆₀ was maintained a few degrees hotter than the iodine to prevent direct

[†] Q.Z., J.E.F., K.K. and O.Z. are also affiliated to the Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6272, USA. Correspondence should be sent to J.E.F.

contact between C_{60} and condensed I_2 and to discourage excess I_2 from crystallizing on the sample as it cooled. The saturation I_2 pressure at 250 °C is 3–4 atm. X-ray powder diffraction profiles measured at beamline X10B of the National Synchrotron Light Source at Brookhaven National Laboratory (wavelength 0.9470 Å) revealed a single, highly crystalline new pattern in all four samples. Substantial amounts of C_{60} (face-centred cubic, $a = 14.17$ Å; ref. 11) were also detected in the samples prepared from $C_{60}:I \approx 1$ and 2, and the $C_{60}:I \approx 12$ sample tube contained large I_2 crystals after reaction. The sample prepared from $C_{60}:I \approx 3$ showed traces only of the strongest C_{60} reflections and no unreacted iodine. Relative intensities of the reflections from the new compound were essentially the same in all four samples. These observations suggest a single stable phase with more than 3 iodine equivalents per C_{60} using the above growth conditions. We also did thermogravimetric analysis (TGA) on 40 mg of the latter sample and found ~40% weight loss after heating to 400 °C, with most of the loss occurring below 200 °C. X-ray analysis of the remaining material revealed only the f.c.c. pattern of pure C_{60} . Attributing the loss to iodine and estimating the effect of small masses of unreacted C_{60} and possible residual iodine, we obtain a composition $C_{60}I_{3.7 \pm 0.7}$.

Figure 1 shows X-ray data taken at 300 K and a fitted profile from the Rietveld refinement described below. The coherence length estimated from the strongest reflection exceeds 1,000 Å. The profile was unchanged after exposing the sample to laboratory air for 2 days. The data were indexed on a simple hexagonal cell with $a = 9.962$ Å and $c = 9.984$ Å. The cell volume per fullerene is 858 Å³, compared with 711 Å³ in f.c.c. C_{60} . Taking 5 Å as the fullerene van der Waals radius, this means that 334 Å³, or 39% of the cell volume, is available for accommodating guest species, as opposed to only 26% in the close-packed f.c.c. structure. Iodine atoms (van der Waals radius ~2.2 Å) can fit easily into trigonal prismatic sites at $\frac{1}{3}, \frac{2}{3}, \frac{1}{2}$ located 7.62 Å away from

six C_{60} centres, or, if crowded slightly, into sites with square planar coordination at $\frac{1}{2}, 0, \frac{1}{2}$ located 7.05 Å away from four C_{60} centres. With an atomic volume of 43 Å³, there is ample room for 4 iodine equivalents per cell.

We compared the results of several Rietveld analyses in order to address the following issues: iodine locations and I–I distances for comparison with bond lengths of neutral I_2 and typical molecular ions such as I_3^- , crystallographic iodine composition for comparison with the TGA and phase mixture results, and possible orientational order of the C_{60} molecules. As there are no systematically absent reflections, the symmetry must be $P6/mmm$ or one of its hexagonal or trigonal subgroups. The former, combined with a spherically averaged C_{60} form factor¹¹, would represent complete orientational disorder, whereas only the trigonal subgroup $P\bar{3}$ is consistent with the molecular point symmetry if complete orientational order is assumed. The diffuse 'background' apparent in Fig. 1 suggests scattering from uncorrelated molecules¹² (at this wavelength, scattering from the glass capillary would peak at $2\theta \approx 16^\circ$). We found no significant difference in the quality of the refinements between ordered and disordered models, no doubt because of the large iodine contribution to the total scattering.

The solid curve in Fig. 1 shows the refined profile assuming orientational order, trigonal space group $P\bar{3}$. This requires 10 C atoms in general positions 6g at x, y, z , the coordinates of which were derived assuming 1.39 Å and 1.45 Å intramolecular C–C bond lengths. The relative orientation of molecular and crystal axes had little effect on the fit. First we fixed I(1) and I(2) atom positions on the high-symmetry 2d and 3f sites $\frac{1}{3}, \frac{2}{3}, \frac{1}{2}$ and $\frac{1}{2}, 0, \frac{1}{2}$ respectively, thus constraining the I–I distances to 3.6 Å, 5.2 Å and 5.9 Å which correspond to neither intra- nor intermolecular distances in solid I_2 (ref. 13). Not surprisingly, this resulted in a mediocre fit and very large iodine thermal factors, the latter usually indicative of large static displacements. Allowing I(1) and I(2) atoms to relax to positions $x, 2x, z$ with a common thermal factor resulted in the very satisfactory fit shown in Fig. 1 ($R_i = 8.24\%$; $R_w = 8.32\%$; thermal factors $B(I) = 4.7$ Å², $B(C) = 2.8$ Å²; refined cell constants $a = 9.9617(3)$ Å, $c = 9.9839(4)$ Å). The refined occupancies are ~2 for each site, corresponding to an ideal formula $C_{60}I_4$, consistent with TGA and phase mixture results.

More significantly, the relaxed I positions yield very reasonable intra- and intermolecular distances. The refined general coordinates of I(1) and I(2) are 0.6423(9), 0.2846(18), $\pm 0.4463(6)$ and 0.5373(4), 0.0745(8), 0.5 respectively (the z coordinate of I(2) was fixed at its ideal position). These would seem to be randomly displaced from the ideal positions by 0.68 Å and 0.64 Å respectively. But locally it is far more plausible that if there is an I(2) atom at 0.5373, 0.0745, 0.5, there will be an I(1) atom on one of the four equivalent sites 0.6423, 0.3577, ± 0.4463 or 0.7154, 0.3577, ± 0.4463 , yielding an I(2)–I(1) distance of 2.53 Å, fairly close to the 2.72-Å intramolecular distance observed¹³ in elemental I_2 . It is then possible to construct a model in which there are two such I_2 molecules per unit cell (one such construction is shown in Fig. 2a), with intermolecular distances of 3.78 and 5.02 Å but with no long-range order. These intermolecular distances compare favourably with the corresponding values 3.50, 3.97 and 4.27 Å observed¹³ in the complex layer structure of solid I_2 . The I(1) and I(2) atoms are situated at distances of 7.12 and 7.08 Å, respectively, from the centres of the C_{60} molecules, and C–I nearest-neighbour separations range from 3.60 to 4.0 Å. Associating an occupied I(1)–I(2) pair with I_2 , the molecules are nearly centred on the trigonal prismatic sites with their axes canted ~11° from the midplane.

This model is not only aesthetically pleasing, but all of the distances are completely consistent with a van der Waals packing of I_2 which preserves many of the local features of the elemental structure. It is also gratifying to note that the composite structure of C_{60} and I_2 sublattices is essentially that of a stage-1 graphite intercalation compound: an eclipsed sequence (AAA...) of host

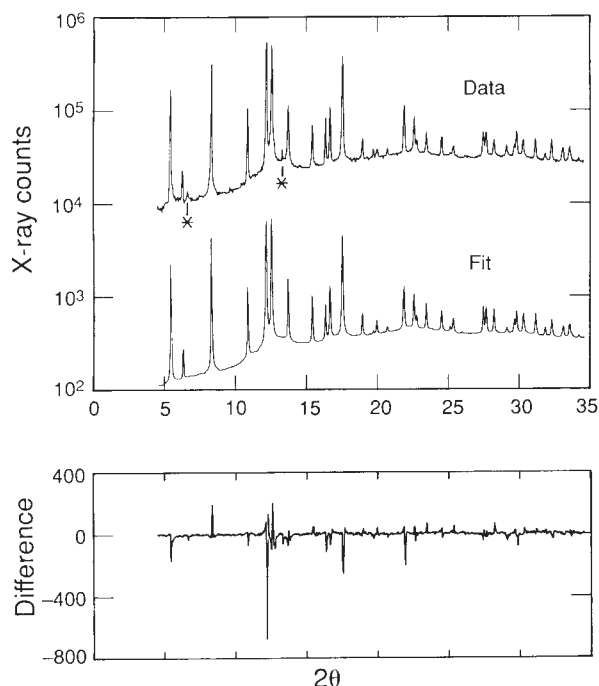


FIG. 1 a, Upper curve: synchrotron X-ray powder diffraction profile of the product resulting from vapour-phase reaction of a 1.5:1 molar ratio of molecular iodine, I_2 , with pure solid C_{60} at 250 °C. Displaced vertically by 2 decades for clarity; note log scale. Wavelength is 0.9470 Å. Weak reflections from unreacted C_{60} are marked by (*). Lower curve: a Rietveld refinement in space group $P\bar{3}$ (trigonal, $a = b = 9.962$ Å, $c = 9.984$ Å and $\gamma = 120^\circ$). The C_{60} reflections were excluded from the refinement. b, Difference between data and fitted profile (linear scale).

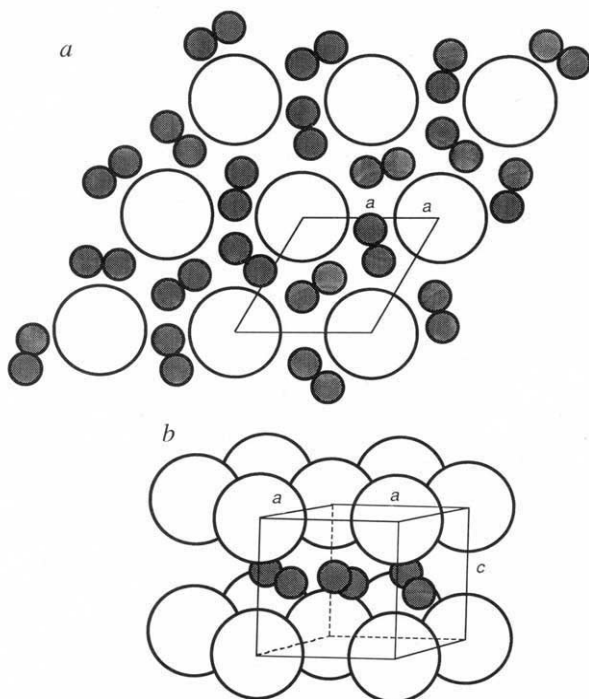


FIG. 2 *a*, Basal plane projection of the refined structure. Large circles are the skeletons of the C₆₀ molecules ($R \approx 3.5$ Å) which form a simple hexagonal lattice, and small circles represent iodine scaled to the van der Waals radius, ~ 2.2 Å. There is no long-range order in the occupancy of iodine sites (see text). *b*, Schematic three-dimensional view of the refined structure using the same scaling. Note the similarity to a stage-1 layer intercalate.

'atom' layers interleaved with guest layers which reside in the van der Waals galleries, as shown schematically in Fig. 2*b*. This suggests that the {111} planes of f.c.c. C₆₀ become the {001} planes in the compound by relative shear motions which transform the layer stacking sequence from ABCABC... to A/A/A/... where '/' denotes a guest layer (the eclipsed configuration being stabilized by guest insertion).

In principle, C₆₀I₄ should show two kinds of order-disorder transitions. One expects *a priori* that a C₆₀ orientational ordering transition will occur significantly above the 249 K value observed¹¹ in pure solid C₆₀, because of the additional steric hindrance from the I₂. Additionally, it is possible that at low temperature the I site occupancy could develop long-range order, probably accompanied by a marked lowering of the symmetry.

The refinement points strongly to molecular (neutral) I₂ as the predominant intercalated species. Saturation doping under similar conditions of a semitransparent C₆₀ film sublimed on mica produced no gross colour changes in transmission or reflection, and no measurable d.c. conductivity. No superconductivity was found down to 4 K. We conclude, in agreement with Ohno *et al.*⁴, that solid C₆₀ is oxidized very little, if at all, by exposure to iodine vapour. But we do find strong evidence for the formation of a definite, possibly nonstoichiometric, highly crystalline compound containing substantial iodine. These different findings are no doubt attributable to our reaction conditions in which the compound is allowed to equilibrate with the saturation vapour pressure of the intercalate. Oxidative intercalation of C₆₀ remains an open question; we are currently searching for a spectroscopic signature of small concentrations of molecular ions such as I₃⁻. For the moment, the new compound C₆₀I₄ seems to be more nearly analogous to the uptake of neutral water by clay minerals than to the ionic salts typified by graphite compounds and the alkali-intercalated fullerenes. On the other hand, C₆₀I₄ does not result from co-crystallization of fullerenes and other molecules (in contrast to the solvated phases

reported by Fleming *et al.*¹⁴) and should therefore be regarded as a true intercalation compound.

Note added in proof: After this paper was accepted, we learned of the publication of very similar work¹⁵. Their experimental X-ray profile, although of much lower resolution, is consistent with the present results, although there are significant differences in the derived structures, in particular regarding the inferred composition. □

Received 2 January; accepted 22 January 1992.

1. Fleming, R. M. *et al.* *Nature* **352**, 701–703 (1991).
2. Zhu, Q. *et al.* *Science* **254**, 545–548 (1991).
3. Stephens, P. W. *et al.* *Nature* **351**, 632–634 (1991).
4. Ohno, T. R., Kroll, G. H., Weaver, J. H., Chibante, L. P. F. & Smalley, R. E. *Nature* (in the press).
5. Swinbanks, T. *Nature* **352**, 749 (1991); **353**, 377 (1991).
6. Lichtenberger, D. L., Nebesny, K. W., Ray, C. W., Huffman, D. R. & Lamb, L. D. *Chem. Phys. Lett.* **176**, 203–208 (1991).
7. Dubois, D., Kadish, K. M., Flanagan, S. & Wilson, L. J. *Am. chem. Soc.* **113**, 7773–7774 (1991).
8. Tsukamoto, J., Takahashi, A. & Kawasaki, K. *J. appl. Phys.* **29**, 125–130 (1990).
9. Fischer, J. E. in *Intercalated Layer Materials* (ed. Levy, F. A.) (Reidel, Dordrecht, 1979).
10. Lefrant, S. *et al.* *Solid State Commun.* **29**, 191–196 (1979).
11. Heiney, P. A. *et al.* *Phys. Rev. Lett.* **66**, 2911–2914 (1991).
12. Neumann, D. A. *et al.* *Phys. Rev. Lett.* (in the press).
13. van Bolhuis, F., Koster, P. B. & Mighelsen, T. *Acta Cryst.* **23**, 90–91 (1967).
14. Fleming, R. M. *et al.* *Phys. Rev. B* **44**, 888–891 (1991).
15. Kobayashi, M. *et al.* *Solid State Commun.* **81**, 93–94 (1992).

ACKNOWLEDGEMENTS. We thank S. Tozer (DuPont) and C. L. Lin (Temple University) for SQUID measurements, and K. S. Liang (Exxon) for the use of the X10B beamline. This work was supported by the NSF Materials Research Laboratory Program and by the Department of Energy. The work at Brookhaven was also supported by DOE, Division of Materials Sciences.

Synthesis of dense silicon-based ceramics at low temperatures

R. Riedel*, G. Passing†‡, H. Schönfelder† & R. J. Brook†‡

* Institut für Anorganische Chemie, Universität Stuttgart, D-7000 Stuttgart 80, Germany

† Max-Planck-Institut für Metallforschung, Institut für Werkstoffwissenschaft, Pulvermetallurgisches Laboratorium, D-7000 Stuttgart 80, Germany

THE conventional preparation of advanced ceramic parts based on silicon carbide or nitride involves pressureless sintering, hot pressing or hot isostatic pressing of appropriate ceramic starting powders¹. Owing to the covalent nature of the Si–C and Si–N bonds and hence the low diffusion coefficients in SiC and Si₃N₄, high sintering temperatures and the addition of sintering aids are normally used to enhance densification. During densification, the sintering additives form second phases located at grain boundaries, which commonly impair the mechanical and physical properties of the material, especially at higher temperatures. New processing routes that overcome these problems are therefore desirable. Here we report the direct transformation of a metallorganic precursor into non-oxide silicon-based ceramics with relative densities of up to 93%. This process can be used to make ceramic components and matrix composites at unusually low temperatures (1,000 °C) and without the addition of sintering aids.

For SiC, common sintering temperatures are 1,800–2,200 °C (ref. 2). Powders of Si₃N₄ have to be sintered between 1,500 and 2,000 °C (ref. 3). An alternative method for the preparation of silicon-based ceramics at reduced temperatures is the pyrolysis of metallorganic compounds, a procedure first applied to the fabrication⁴ of continuous SiC-fibres from silicon polymers such as polycarbosilanes ([–RSiH–CH₂–]_n, where R is CH₃). Pure Si₃N₄ or mixtures of SiC and Si₃N₄ ceramic powders and fibres have been similarly obtained by the thermal decomposition of organo-substituted polysilazanes ([–RSiR'–NH–]_n, where R is

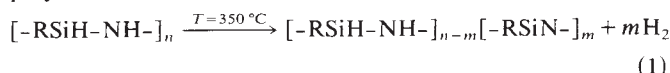
‡ Present addresses: Bayer AG, D-5090 Leverkusen, Germany (G.P.); and Department of Materials, University of Oxford (R.J.B.).

H or CH₃ and R' is CH₃) in ammonia or in protective atmospheres of N₂ or Ar respectively⁵⁻⁷. Nevertheless, silicon-based polymers cannot be directly transformed into larger-scale ceramic parts by this process because the release of gaseous by-products during the mineralization step causes foaming and crack formation. We have found, however, that dense, crack-free Si₃N₄ and Si_xN_yC (silicon carbonitride) ceramics can be obtained at temperatures as low as 1,000 °C if an infusible polymethylsilazane powder is used as the starting material. In Fig. 1 our polymer process route is summarized schematically.

Infusible polymethylsilazane was synthesized by heat-treating commercial polysilazane NCP 200 (Nichimen Corporation, Japan) at 350 °C for 100 min in flowing argon atmosphere and finally by distillation of volatiles in vacuum (2×10^{-3} Torr) for 80 min. Subsequently, the annealed polymeric product was crushed and ball-milled, isostatically pressed (640 MPa) to form green compacts (60 × 22 × 14 mm³) or cylindrical samples (1 cm diameter and 1 cm height) and finally pyrolysed in quartz tubes under argon or ammonia at 1,000 °C using a heating rate of 5 K min⁻¹. Thermal decomposition was studied by mass spectroscopy of the gaseous compounds evolved during heat treatment in the temperature range between room temperature and 1,000 °C. We analysed the ceramic residue decomposed in Ar for its C, N, O and Si content. We measured pore size distributions and densities of the green and pyrolysed samples by mercury pressure porosimetry, and also measured densities using the Archimedes method. The mechanical strength (σ_B) was determined at room temperature in four-point bend tests of silicon carbonitride rods (2.2 × 2.3 × 42 mm³). We subsequently heat-treated the pyrolysed samples at 1,200–1,600 °C (graphite furnace) in Ar and N₂ to investigate the crystallization behaviour and to determine the weight change and dimensional changes of the samples. The oxidation was studied by thermal gravimetric analysis (TGA) of the pyrolysed compacts at 1,200–1,500 °C in air.

The starting polymeric compound, polysilazane NCP 200, contains [–RSiH–NH–] and [–R₂Si–NH–] units, where R is CH₃ (information provided by Nichimen Co., Tokyo). Infrared spectroscopic investigations have confirmed that no aromatic

or vinylic groups are present in the precursor. At temperatures exceeding 200 °C, the reaction of Si–H and N–H groups takes place with the simultaneous loss of NH₃ and H₂ leading to the formation of trisilylated nitrogen bridges (NSi₃) between the polymeric molecules:



According to mass spectroscopic measurements, hydrogen ($m/e=2$) and ammonia ($m/e=17$) are evolved in the temperature range 250–350 °C. During isothermal heat-treatment of as-received polymethylsilazane NCP 200 at 350 °C in flowing Ar for 3 or 6 h, we have observed mass losses of 1.8 and 3.0 wt%, respectively. This, together with the occurrence of transamination reactions, is assumed to cause the high degree of crosslinking and results in the conversion of the tractable polymethylsilazane into an unmeltable and hence insoluble polymeric silazane. Infusibility is necessary for the shape of compacted polymeric powder samples to be retained during pyrolysis. In our process it is important therefore that the chemical nature of the polymer allows such crosslinking, either thermally or chemically.

In a second step, the infusible polymer is crushed or milled and subsequently formed into shapes by uniaxial or cold-isostatic pressing (Fig. 2a). During pressing, the polymethylsilazane compacts achieve relative densities of ~84% (we take 1.22 g cm⁻³ as the solid phase density from pycnometer measurements). The high densities of the shaped polymer indicate viscous flow in the polymethylsilazane during pressing. The final step involves the thermal decomposition of the crosslinked polymethylsilazane. The main gaseous decomposition products evolved are CH₄ (400–750 °C; maximum at 525 °C) and H₂ (150–900 °C; maxima at 500 and 700 °C) as determined by TGA–MS measurements. According to thermal gravimetric analysis, the total weight loss up to 1,000 °C is 24% and the pyrolysis reaction is nearly complete at 800 °C, the weight loss between 800 and 1,000 °C is ≤1 wt%.

Owing to the open porosity present in the compacted polymeric green samples, the gaseous reaction products can be

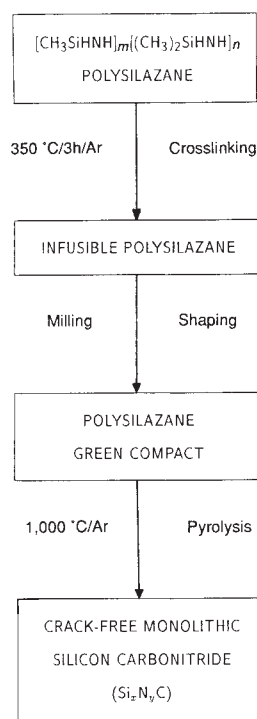


FIG. 1 Schematic representation of the silazane into dense and crack-free low-temperature process route for the silicon carbonitride or silicon nitride direct conversion of polymethyl- compacts.

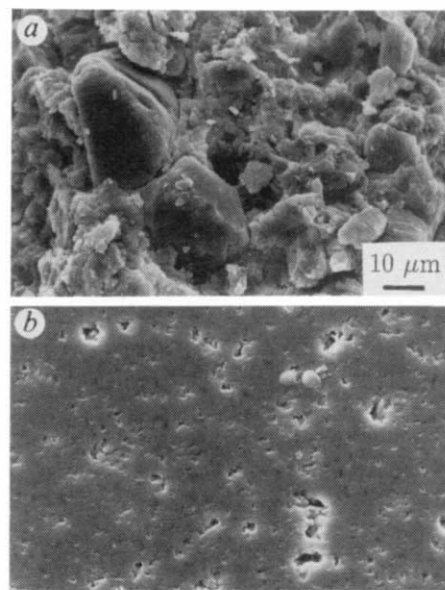


FIG. 2 Scanning electron micrographs of a, a fracture surface of infusible polymethylsilazane powder isostatically pressed at 640 MPa ($\rho=84\%$ relative density) and b, a polished surface of dense and crack-free silicon carbonitride ($\rho=93\%$ relative density) derived from infusible polymer compacts heat-treated at 1,000 °C in Ar.

removed without the formation of cracks. The original polymer powder is converted into an inorganic solid and shrinks linearly by 20–25% with simultaneous change in density. The resulting ceramic parts are X-ray amorphous and crack-free, and have relative densities up to 93%; they have good wear resistance and cannot be scratched by steel. The good mechanical resistance itself suggests the formation of new covalent bonds between the individual particles. The microstructure of the resulting continuous three-dimensional network is shown in the scanning electron micrograph of Fig. 2b. No macroscopic cracks can be detected, and pore sizes over 10 μm are not found in the polymer-derived ceramic monoliths. At higher magnification, the microscope reveals dense and pore-free particles bonded together without a second phase in the microstructure. Analytical transmission electron microscopy shows a homogeneous distribution of carbon and nitrogen in the range of ± 20 Å; no segregated particles in the form of amorphous C, SiC or Si_3N_4 can be detected in the amorphous silicon carbonitride.

In contrast, heat-treatment of polysilazane NCP 200 at 350 °C for 6 h yields a crosslinked but unreactive polymeric powder which cannot then be transformed into wear-resistant parts. Thus, the presence of reactive polymer particles is necessary to obtain the interconnection of the polymeric granules during the mineralization step.

The black polymethylsilazane-derived compacts pyrolysed at 1,000 °C in argon contain 57.8 wt% Si, 14.3 wt% C and 26 wt% N, corresponding to an amorphous silicon carbonitride with the composition $\text{Si}_{1.73}\text{N}_{1.56}\text{C}_{1.0}$; the oxygen contamination is determined to be below 0.8 wt%. If ammonia is used as a reactive gas, pure colourless Si_3N_4 can be obtained by the pyrolysis of polysilazanes⁸.

Mercury pressure porosimetry of the silicon carbonitride parts reveals an absolute density of 2.15 g cm^{-3} , corresponding to 93% relative density on the basis of the solid phase density of 2.30 g cm^{-3} found in pycnometric measurement of the powdered pyrolytic residues. The pore size distribution shown in Fig. 3 has maximum pore radii up to 110 nm, indicating that the larger pore sizes seen in the electron micrograph (Fig. 2b) stem from isolated pressing faults generated during the shaping of the crosslinked polymethylsilazane pieces. In contrast to the compacted polymer powder samples, specimens heat-treated at 1,000 °C in Ar show a narrower pore size distribution in mercury pressure porosimetry (Fig. 3). All open pores less than 15 nm in radius have been removed. Furthermore, the cumulative pore volume has been decreased owing to densification induced by the loss of volatiles and the simultaneous formation of the ceramic phase.

The pyrolytic compacts currently have mechanical strengths

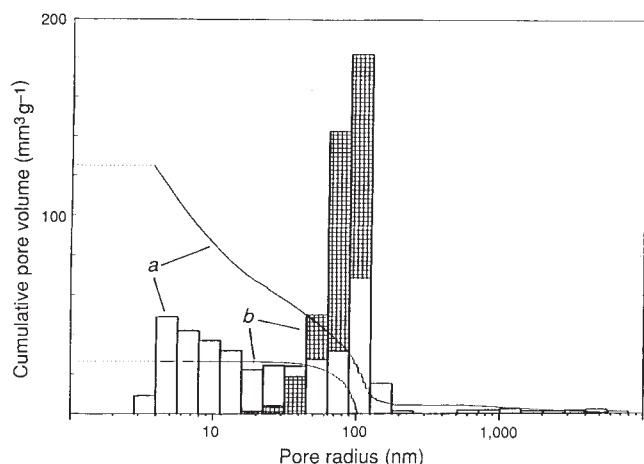


FIG. 3 Pore size distribution of infusible polymethylsilazane, a, isostatically compacted at 640 MPa and b, pyrolysed at 1,000 °C in Ar.

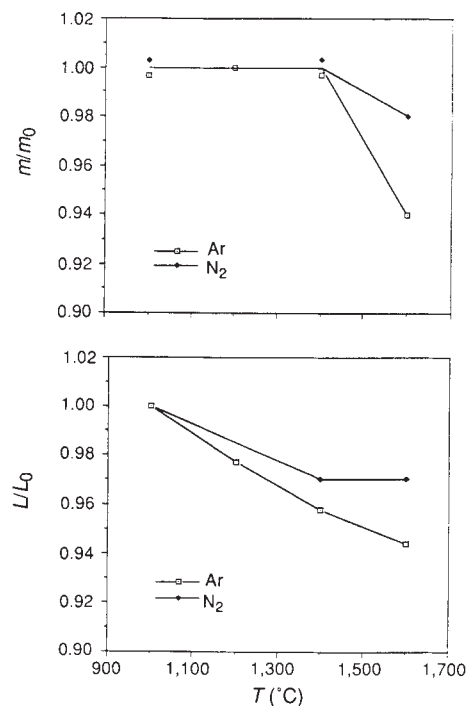


FIG. 4 Relative change in mass (m/m_0) and change in length (L/L_0) of silicon carbonitride ($\text{Si}_{1.73}\text{N}_{1.56}\text{C}_{1.0}$) during heat-treatment in the temperature range 1,000–1,600 °C in Ar and N_2 atmosphere.

up to 375 MPa. The Vickers hardness (9.5 GPa) is low in comparison to the hardnesses of pure and crystalline SiC (27–30 GPa) and Si_3N_4 (18–20 GPa). Reaction-bonded silicon nitride (RBSN) with a porosity similar to the present material has a hardness of 8.9 GPa (ref. 9).

The amorphous silicon carbonitride structure can be retained up to 1,400 °C in Ar and up to 1,600 °C in a N_2 atmosphere. In Ar, SiC begins to crystallize in the β polymorph at 1,400 °C, whereas in N_2 , α - Si_3N_4 as well as α - and β -SiC crystals are first formed at temperatures above 1,600 °C. In contrast, amorphous Si_3N_4 derived from pyrolysis of polymethylsilazane (NCP 200) in NH_3 starts to crystallize at 1,000 °C to α - Si_3N_4 , as could be confirmed by X-ray diffraction. Silicon carbonitride samples heat-treated for 1 h in N_2 at 1,600 °C shrink linearly by 3%. Those treated in Ar shrink by 5.5%, consistent with the crystallization. The relative change in mass (m/m_0) and change in length (L/L_0) in the temperature range between 1,000 and 1,600 °C can be taken from Fig. 4.

Investigation of the oxidation behaviour of the silicon carbonitride compacts with 8% porosity reveals that no weight change occurs within 24 h at 1,300 °C in air. At 1,400 °C, oxidation results in a weight gain of 4%, whereas the mass of a sample annealed at 1,500 °C in air increases by 7%. The oxidation rate decays rapidly with time in both cases so that most of the mass gain is achieved in the first hour of the isothermal soak; after 24 h at 1,500 °C the rate falls to considerably less than $0.1\% \text{ h}^{-1}$.

The microstructure of our silicon carbonitride materials can be designed to be amorphous or nanocrystalline, as shown previously^{8,10}. Physical and mechanical properties different from those of conventionally processed SiC- and Si_3N_4 -ceramics can be expected as a consequence of the low processing temperature and the stability of the compositionally complex amorphous phases. The stability is analogous to the invert glasses of Stevels¹¹ but is found here at a much higher temperature range and in nonoxide systems. In addition, the energy costs associated with the densification of crystalline ceramic powders at high temperatures (SiC, for example, requires a temperature of 2,100 °C) are avoided. □

Received 25 September; accepted 28 November 1991.

- Greil, P., Petzow, G. & Tanaka, H. *Ceram. Int.* **13**, 19–25 (1987).
- Böcker, W., Landfermann, H. & Hausner, H. *Powder Metallurgy Int.* **11**, 83–85 (1979).
- Ziegler, G., Heinrich, J. & Wötting, G. *J. Mat. Sci.* **22**, 3041–3086 (1987).
- Yajima, S., Hayashi, J., Omori, M. & Okamura, K. *Nature* **261**, 683–685 (1976).
- Seyferth, D. & Wiseman, G. *J. Am. Ceram. Soc.* **67**, C132–133 (1984).
- Riedel, R., Seher, M. & Becker, G. *J. Eur. Ceram. Soc.* **5**, 113–122 (1989).
- Peuckert, M., Vaahs, T. & Brück, M. *Adv. Mater.* **2**, 398–404 (1990).
- Riedel, R. & Seher, M. *J. Eur. Ceram. Soc.* **7**, 21–25 (1991).
- Ritter, J., Nair, S., Gennon, P. & Dunlay, W. *Adv. Ceram. Mater.* **3**, 415–417 (1988).
- Riedel, R., Strecker, K. & Petzow, G. *J. Am. Ceram. Soc.* **72**, 2071–2077 (1989).
- Stevens, J. M. in *Handbuch der Physik* **20**, 350 (Springer, Berlin, 1957).

Reduction of nitrate and nitrite in water by immobilized enzymes

Robert B. Mellor, Jörg Ronnenberg,
Wilbur H. Campbell* & Stephan Diekmann

Mobitec GmbH, Wagenstieg 5, D-3400 Göttingen, Germany

* Phytotechnology Research Center, Michigan Technological University,
Houghton, Michigan 49931, USA

NITRATE, a common and serious contaminant of ground water, is removed at present either by physicochemical methods that do not degrade it, or via degradation by microorganisms, which is a slow process¹. We report here a rapid and efficient process for nitrate removal which involves catalytic reduction by immobilized enzymes. The reduction is driven by an electrical current, and results in complete conversion of nitrate to N₂ without residues. Our electro-bioreactor was constructed by co-immobilizing the enzymes (purified NADH:nitrate reductase from *Zea mays*² and crude nitrite reductase and N₂O reductase from *Rhodopseudomonas*³) with electron-carrying dyes in a polymer matrix, which was then attached in thin layers to the surface of the cathode. Nitrate-laden water is pumped past the anode and through the active matrix on the cathode while a low voltage is applied, resulting in two-stage nitrate reduction to N₂ via nitrite. The enzyme activity is higher in the co-immobilized state than in free solution. In principle, such electro-bioreactors could be developed for removal of other water contaminants such as pesticides, if appropriate enzymes and cofactors can be identified.

The widespread pollution of ground water with nitrate in Europe and the United States is an environmental problem and a health hazard⁴. In many regions, the nitrate concentration is at or near the accepted standards for potable water, 40–50 mg of nitrate per litre⁵. Existing methods for removing nitrates are not effective, and often the contaminated water source is simply abandoned, especially in the case of wells for individual residences. Recently, an alternative was suggested where the nitrate was removed chemically, but a residual aluminium contamination was introduced in the process⁶. Biogenic denitrification using microorganisms requires that the water be enriched with nutrients to support their growth^{1,7,8}. Furthermore, these microorganisms are difficult to maintain, and the overall process of denitrification is both slow and often incomplete, producing nitrous oxides which then become air pollutants. Here we use only the denitrification enzymes in a bioreactor instead of using whole bacterial cultures.

To construct a bioreactor, one needs readily available sources of the denitrification enzymes. Although many of the enzymes involved in denitrification have been isolated and characterized, some are membrane-bound and therefore difficult to manage *in vitro*^{9,10}. Soluble forms of nitrite reductase and nitrous oxide reductase have been isolated and characterized from a number of microorganisms, and these would be suitable for a bioreactor catalysing denitrification^{3,7–10}. But soluble forms of nitrate reductase have been described only occasionally from bacterial sources, and it is generally accepted that the membrane-bound nitrate reductase is the enzyme involved in denitrification⁹. On

TABLE 1 Activities of nitrate reductase, nitrite reductase and nitrous oxide reductase after immobilization

Reductase	Immobilization pH	Electron donor	Relative activity (%)
Nitrate	7	methyl viologen	43.9
		NADH	1.3
Nitrate	4	methyl viologen	43.4
		NADH	1.6
Nitrite	7	methyl viologen	89.5
Nitrite	4	methyl viologen	81.9
N ₂ O	7	methyl viologen	77.8
N ₂ O	4	methyl viologen	69.9

Relative activities were the mean of four independent determinations. 0.1 unit of activity for each enzyme was separately immobilized on 200 µl of vinyl polymer matrix which had been modified by the method of ref. 17. When immobilization was done at 22 °C, all three enzymes were completely bound to the gel. NADH:nitrate reductase (EC 1.6.6.1) from *Zea mays* was purified by immunoaffinity chromatography and assayed as previously described (1 unit = 1 µmol nitrite formed per min)². Dissimilatory nitrite reductase (EC 1.7.2.1) and nitrous oxide reductase (no EC number) were isolated together from the periplasmic space of *Rhodopseudomonas spheroides* forma *denitrificans* (gift of Professor Satoh, University of Tokyo) and assayed as previously described (1 unit = 1 µmol nitrite removed per min and 1 µmol nitrogen gas produced per min, respectively)³. All enzyme assays were done at 22 °C. Immobilized enzymes were assayed with the substrates indicated. For nitrate and nitrite reductase, the matrix was washed with 2.5 ml buffer to recover nitrite and determine its amount. For nitrous oxide reductase, the assay was carried out in a 3-ml tube fitted with a rubber septum. The reaction was initiated by injection of 100 nmol nitrous oxide into helium atmosphere over the matrix. The decrease in nitrous oxide and increase in nitrogen in the gas phase were measured by gas chromatography¹⁸.

the other hand, the assimilatory nitrate reductase of eukaryotes is a soluble enzyme and, being a highly effective catalyst, it would be suitable for use in a denitrifying bioreactor when coupled with the bacterial denitrifying enzymes¹¹. Recently, purified forms of nitrate reductase from higher plants have become available which are stable *in vitro*². We have therefore developed a method for nitrate removal from water using assimilatory NADH:nitrate reductase from *Zea mays* in combination with nitrite reductase and nitrous oxide reductase, which are soluble enzymes from the periplasmic space of *Rhodopseudomonas spheroides*³.

TABLE 2 Activity of reductases with immobilized dyes and when co-immobilized with the same dyes

Immobilized dye	Nitrate reductase		Nitrite/N ₂ O reductases	
	free	bound	free	bound
Azur A	30.8	97.6	52.7	62.5
Safranin T	44.5	75.0	55.7	104.1
Neutral red	55.0	16.4	68.2	102.9
Bromophenol blue	21.5	93.9	58.3	50.0
Cibacron blue	38.6	20.4	88.2	55.3

Activities of the enzymes are shown as a percentage of their activity with immobilized methyl viologen. The values are the mean of four independent assays. An error of ±10% was found. For the assay of free enzyme, 0.1 unit of activity was mixed with the immobilized dyes, whereas for the assay of bound enzyme, 0.1 unit activity was co-immobilized with the immobilized dyes. The assay method was as described in Table 1, except that only nitrite decrease was assayed for the mixed nitrite/nitrous oxide reductases. Dyes were from Fluka, except Cibacron blue 3GA and bromophenol blue, which were from Sigma. Dyes containing free amino groups were coupled at pH 7.0 in aqueous solution for 30 min at room temperature for full binding and for 10 min when co-immobilized with enzymes. Bromophenol blue with no free amino group was coupled in dry acetone for 2 h (co-immobilization) or 8 h (full binding).

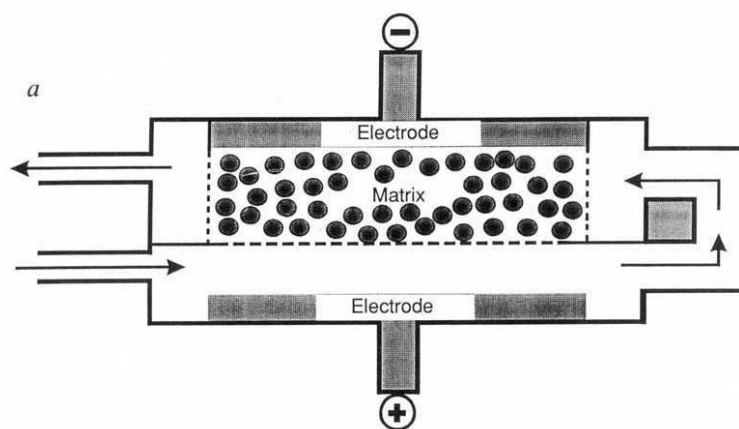
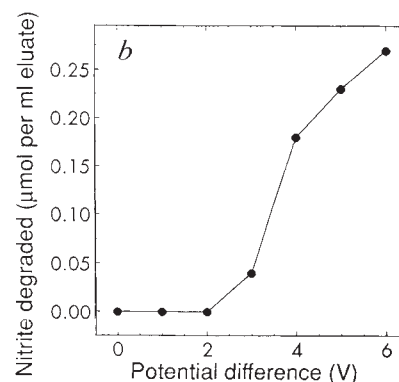


FIG. 1 *a*, Schematic representation of an electrical column (electrodes 9.5 cm × 4.5 cm conductive plate) with 2 ml nitrite reductase/ N_2O reductase/Safranin T matrix (2 units enzyme) packed within 1 mm of the cathode. Arrows indicate the direction of water flow. Dashed line indicates the position



of a 90-μm pore-size glass filter. *b*, Nitrite reductase activity in the above column with increasing potential difference across the cell. Flow rate: 10 ml min⁻¹. Substrate concentration, 1 mM.

We developed a bioreactor for nitrate removal from water in two phases: first, establishing a method of immobilizing the enzymes in such a way that they retain their activities, and second, developing an effective method to deliver reducing power to the immobilized enzymes. Plant nitrate reductases use pyridine nucleotides as the natural electron donor, but can accept electrons from artificial donors such as the viologen dyes¹² and also work very effectively with bromophenol blue as the electron donor¹³. When we immobilized maize nitrate reductase, it lost virtually all its activity with NADH, but retained more than 40% of its nitrate-reducing activity with methyl viologen (Table 1). We found that bacterial nitrite reductase and nitrous oxide reductase were also highly active when immobilized, with methyl viologen as the electron donor. Because the toxicity of methyl viologen makes it unacceptable for purifying drinking water¹⁴, we analysed other dyes as electron donors for the immobilized enzymes and found several equally effective substitutes (Table 2). To optimize the functionality of the immobilized enzymes in a bioreactor, we developed a method for co-immobilizing the electron-carrying dyes with the enzymes. In the course of this, we found that dyes that were poor electron donors for the immobilized enzymes when transferring electrons from solution were much more effective when co-immobilized with the enzymes (Table 2). The opposite was found for dyes that were highly effective in solution, suggesting that the enzymes interacted with the bound dyes during the co-immobilization process. We concluded that the enzymes were perhaps binding to the immobilized dyes in an orientation favourable for electron

transfer and then becoming covalently attached to the matrix beads, and that the overall process was optimized when the enzyme had a high affinity for the oxidized dye. On this basis, we co-immobilized the enzymes with dyes under different external conditions and were able to increase the activity of the system to values which are 30% higher than the activity in solution (data not shown).

The problem of effectively delivering reducing equivalents to the co-immobilized enzymes and electron-carrying dyes was solved by designing an electro-bioreactor where the reducing power was provided by an electrical current. We found that the most effective reduction of nitrate was achieved when the matrix containing the enzymes was entrapped in a thin layer over a large cathode surface, and the water first flowed past the anode (Fig. 1*a*). Low applied voltages effectively degraded nitrite (Fig. 1*b*). Above 1.28 V, the electrochemical potential of water, water molecules are effectively protonated at the anode. These hydroxonium ions flow to the cathode, receive an electron, and decompose to water and atomic hydrogen. Our experiments showed that redox indicator dyes could effectively capture such atomic hydrogen before it joined to molecular dihydrogen or at least mediate the flow of current in the electro-bioreactor, resulting in an effective transfer of electrons to the redox enzymes co-immobilized near the reduced dyes. In the final electro-bioreactor, two stages were used: the first for nitrate reduction to nitrite, where a matrix of co-immobilized nitrate reductase and azur A was entrapped over the cathode, and the second for nitrite reduction to N_2 , where a matrix of co-immobilized nitrite

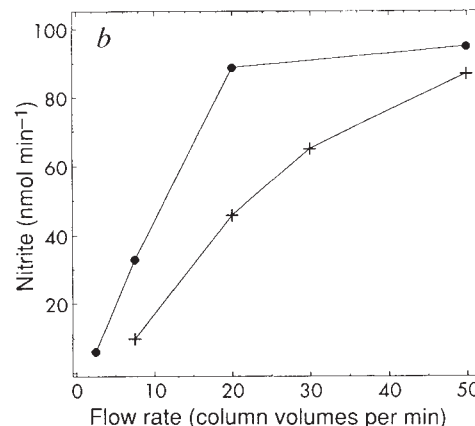
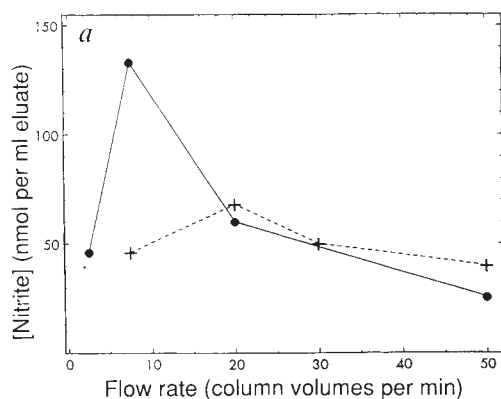


FIG. 2 Nitrate (●) or nitrite (+) turnover in an electrical column as in Fig. 1 (3 V, packed with one unit either nitrate reductase/Azur A or nitrite reduc-

tase/Safranin T matrix) against flow rate, expressed either as (a) turnover per ml or (b) total efficiency. Substrate concentration, 1 mM.

reductase/nitrous oxide reductase and safranin T was entrapped over the cathode. We established the effectiveness of both stages by monitoring the appearance of nitrite after the first stage and its disappearance after the second stage at different flow rates (Fig. 2). Furthermore, analysis of the stoichiometry of the nitrate- and nitrite-reducing system indicated that when the two reactor stages were coupled together, nitrate was completely converted to N_2 without production of nitrous oxide. From our results we calculated the specific catalytic activity of the electro-bioreactor to be at least 560 kg nitrate per m^3 of matrix per day. Preliminary tests of the system indicate that the electro-bioreactor was stable over a period of 3 months, retaining more than 50% of its original activity. More intensive field tests are required to establish the long-term stability of the system when in continuous operation.

In principle, electro-bioreactors could be developed for any water pollutant for which an appropriate redox enzyme exists. For example, pesticides can be degraded by cytochrome P-450 systems¹⁵, which are redox enzymes with some similarity to those used here to degrade nitrate. Indeed, some soluble forms of cytochrome P-450 have been described¹⁶ which might be suitable for applications in electro-bioreactors.

Oxidoreductases represent the largest class of known enzymes, so a large range of substrates could potentially be degraded in

electro-bioreactors. Their full technological power has not yet been realized, possibly because these enzymes are usually used in their soluble form. When immobilized, they could become widely applicable and powerful tools in environmental biotechnology. □

Received 2 September; accepted 4 December 1991.

1. Weil, G. & Kalbertodt, S. *Bioengineering* **6**, 21–31 (1989).
2. Hyde, G. et al. *Plant molec. Biol.* **13**, 233–246 (1989).
3. Neu, H. D. & Heppel, L. A. J. *Biol. Chem.* **240**, 3685–3692 (1965).
4. Mirvish, S. *Nature* **315**, 461–462 (1985).
5. *Eur. Chem. Industry Ecol. Toxicol. Centre Tech. Rep.* 27, Brussels (1988).
6. Murphy, A. P. *Nature* **350**, 223–225 (1991).
7. Michalski, W. P. & Nicholas, D. J. D. *Biochim. biophys. Acta* **828**, 130–137 (1985).
8. Michalski, W. P. et al. *Biochim. biophys. Acta* **872**, 50–60 (1986).
9. Hochstein, L. I. & Tomlinson, G. A. *Ann. Rev. Microbiol.* **42**, 231–261 (1988).
10. Ferguson, S. J. *Trends biochem. Sci.* **12**, 354–357 (1987).
11. Campbell, W. H. *Physiol. Plant.* **74**, 214–219 (1988).
12. Barber, M. J. & Solomonson, L. P. *J. Biol. Chem.* **261**, 4562–4567 (1986).
13. Campbell, W. H. *Plant Physiol.* **82**, 729–732 (1986).
14. *Registry of Toxic Effects of Chemical Substances, Natl. Inst. Occupational Safety and Health* DW 2275000 (Bethesda, 1982).
15. Black, S. D. & Coon, M. J. *Adv. Enzym.* **60**, 35–87 (1987).
16. Porter, T. D. *Trends biochem. Sci.* **16**, 154–158 (1991).
17. Biebricher, C. & Luce, R. *US Patent No.* 4, 177,038 (1979).
18. Urata, K., Shimada, K. & Satoh, T. *Plant Cell Physiol.* **23**, 1121–1124 (1982).

ACKNOWLEDGEMENTS. W.H.C. thanks the NSF for financial support. This work is supported by the German Ministry for Research and Technology by a grant to Mobitec within the TOU program.

Earliest *Homo*

Andrew Hill*, Steven Ward†, Alan Deino‡, Garniss Curtis‡ & Robert Drake‡

* Department of Anthropology, Yale University, New Haven, Connecticut 06520, USA

† Department of Human Anatomy, Northeastern Ohio Universities College of Medicine, Rootstown, Ohio 44272, USA

‡ Geochronology Centre of the Institute of Human Origins, 2453 Ridge Road, Berkeley, California 94709, USA

THE origin of our own genus, *Homo*, has been tentatively correlated with worldwide climatic cooling documented at about 2.4 Myr (million years) (refs 1–5). It has also been conjectured that members of *Homo* made the first stone tools, currently dated at 2.6–2.4 Myr (refs 6–8). But fossil specimens clearly attributable to *Homo* before about 1.9 Myr have been lacking. In 1967 a fossil hominoid temporal bone (KNM-BC1) from the Chemeron Formation of Kenya was described as family Hominidae gen. et sp. indet.⁹. Although a surface find, its provenance within site JM85 (BPRP site K002) was established and a stratigraphic section provided indicating the specimen's position⁹. This evidence has been affirmed (see for example refs 10–12) but the exact age of the fossil was never determined, and the absence of suitable comparative hominid material has precluded a more definitive taxonomic assignment. Here we present ⁴⁰Ar/³⁹Ar age determinations on material from the hominid site indicating an age of 2.4 Myr. In addition, comparative studies allow us to assign KNM-BC1 to the genus *Homo*, making it the earliest securely known fossil of our own genus found so far.

As described by Tobias⁹, KNM-BC1 is an undistorted partial right hominid temporal lacking portions of its squama, mastoid and zygomatic processes but sufficiently complete for detailed comparison with other hominid temporal bones that have been shown to possess many important diagnostic elements^{13–17}. We have determined that the Chemeron temporal shares the following features with *Homo*: a medially positioned mandibular fossa that is mostly situated beneath the middle cranial fossa, and a sharp petrous crest. These features collectively suggest possible differences in the functional anatomy of the temporomandibular joint when compared with *Australopithecus*, as well as possibly providing evidence for brain expansion¹⁵ in early *Homo*. Two features present in the Chemeron temporal bone, a vertical and

deep tympanic plate, and a distinct, sharp tympanic crest, appear to be present in specimens attributable to early *Homo*, *A. africanus* and *A. robustus*.

KNM-BC1 differs from the African apes in possessing a deep, medially situated temporomandibular joint fossa, minimal inflation of the temporal squama by mastoid air cells, and sharp tympanic crest. It differs from the described temporal anatomy of *A. afarensis* for almost precisely the same reasons it differs from the African apes, underscoring the probably primitive anatomical pattern diagnostic of the earliest known hominid species (Table 1; Fig. 1a). The Chemeron temporal is also distinct from *A. boisei*, which has a laterally placed mandibular fossa and a massive articular tubercle. The tympanic plate is shallow, and the tympanic crest is expressed as a blunt ridge. The anterior margin of the joint surface is sharply delimited, ending abruptly at the temporal fossa. Despite these differences, *Homo*, *A. robustus* and *A. boisei* share several presumably convergent features, including a steep posterior slope of the articular eminence, vertical tympanic plate, and reduction of the postglenoid process¹⁸. It is likely that these and perhaps other elements of basicranial topography shared by these taxa are secondary to pronounced basicranial flexion, which appears to have developed independently in the *Homo* and robust australopithecine lineages. KNM-BC1 differs from KNM-WT 17000 for similar reasons, although there are important differences in temporomandibular joint and tympanic morphology that distinguish KNM-WT 17000 from *A. boisei*¹⁹. In addition, KNM-WT 17000 has a fully inflated squamous temporal, a feature it shares with *A. afarensis*²⁰.

The features that appear to be unique to members of the *Homo* clade (medial temporomandibular joint, and sharp petrous crest for attachment of the tentorium cerebelli) in conjunction with features in which it differs from KNM-WT 17000, *A. afarensis*, *A. robustus*, *A. boisei* and *A. africanus* (Table 1) mandate attribution of KNM-BC1 to *Homo* sp. indet. Comparison with the condition found in African apes (Table 1) suggests that these are not simply shared primitive features of Hominoidea, with hominids other than *Homo* being derived in these respects. Moreover, several features, such as shape of the external acoustic meatus and pattern of supramastoid and nuchal crest inflation, indicate that the specimen is not assignable to either *H. erectus* or *H. sapiens* (Table 1).

We have measured ⁴⁰Ar/³⁹Ar ages from anorthoclase separated from pumice lapilli (RD-6) and a 20-cm pumice block

TABLE 1 Distribution of hominid temporal bone features

	African apes	<i>A. afarensis</i>	KNM WT 17000	<i>A. africanus</i>	<i>A. robustus</i>	<i>A. boisei</i>	<i>H. habilis</i>	<i>H. erectus</i>	<i>H. sapiens</i>	KNM BC1
Squamous temporal pneumatized	×	×	×	—*	—	—	—	—	—	—
Enlarged temporomandibular tubercle	—	—	—	—	×	×	—	—	—	—
Tympanic plate vertical, deep	—	—	×	×	×	×	×	×	×	×
Deep temporomandibular fossa	—	—	—	—	—	—	×	×	×	×
Tympanic crest sharp	—	—	—	×	×	—	×	×	×	×
Mandibular fossa medial	—	—	—	—	—	—	×	×	×	×
Extended preglenoid surface, TMJ	—	×	—	×	—	—	×	×	×	×
Sharp petrous crest	—	—	—	—	—	—	×	×	×	×
Supramastoid crest inflated	×	×	×	×	×	×	×†	×	—	×
Circular external auditory meatus	—	×	×	×	×	×	×	—	—	×

Squamous temporal pneumatization as identified by Kimbel¹⁴ occurs only in *A. afarensis* and KNM-WT 17000 among hominids, although some *A. africanus* specimens (Sts 19, MLD 37/38 and at least one *A. boisei* (KNM-WT 17400)) show a tendency towards squamous temporal inflation. A similar, and presumably developmentally homologous pattern is present in living pongids. A more circumscribed incursion of mastoid cells into the temporal squama can occur in all other known hominids, but clearly represents a derived pattern. Enlarged temporomandibular tubercle at the lateral margin of the articular eminence is the point of origin of the temporomandibular ligament and unambiguously distinguishes *A. boisei* and *A. robustus* from all other hominid species, including KNM-WT 17000. *A. boisei* has a massively expanded zygomatic root, with a massive temporomandibular tubercle. Other species within the robust clade show evidence of lateral zygomatic expansion, presumably reflecting enlargement of the posterior parts of the temporalis muscle. It is also possible that the lack of a pronounced temporomandibular tubercle in KNM-WT 17000 is associated with its more open or unflexed cranial base when compared with *A. boisei* crania³⁶. A vertical tympanic plate characterizes all known species of *Homo* as well as *A. boisei* and *A. robustus*. A deep temporomandibular fossa in which the mandibular condyle and the meniscus of the temporomandibular joint are separated from the middle cranial fossa by a thin plate of bone appears to occur only in *Homo*. An apparently convergent pattern in *A. robustus* and *A. boisei* in which the fossa appears deep is in fact mediated by the enlarged articular eminence and temporomandibular tubercle. A sharp tympanic crest occurs in all species of *Homo*, *A. africanus*, *A. robustus* and *A. boisei* in contrast to *A. afarensis* in which the tympanic is more tubular. When present a distinct tympanic crest terminates medially in a variably expanded and flange-like vaginal process for the styloid. Medial position of the mandibular fossa is characteristic of *Homo*. This topographic arrangement reflects significant changes in the relationships of the cranial base and is a consequence of cerebral expansion, and to a lesser extent reduction in the size of the zygomatic root. An extended preglenoid surface onto the infratemporal surface of the sphenoid occurs in all hominids other than *A. boisei* and *A. robustus*. Although broken at the zygomatic root, KNM-BC1 clearly shows the non-'robust' configuration of limited anterior exposure of the TMJ to the temporal fossa. A sharp petrous crest appears to be present only in some species of *Homo* and may be associated with expansion of the supratentorial part of the cerebrum¹⁵. Inflation of the supramastoid crest occurs in some specimens attributed to early *Homo* (KNM-ER 1805, KNM-ER 1470, KNM-ER 1813, among others) as well as *A. afarensis* and KNM-WT 17000. But it is possible that supramastoid inflation is not strictly homologous in *A. afarensis* and early *Homo*. Circular external auditory meatus segregates *Australopithecus* and early species of *Homo* from *H. erectus* and *H. sapiens*. In the former, the external meatus is essentially circular, and in later species of *Homo* it tends to be antero-posteriorly compressed, with its greatest dimension in the vertical axis.

* At least two *A. africanus* crania, Sts. 19, and MLD 37/38 show evidence of squamous temporal pneumatization.

† Some specimens attributable to early *Homo* by some workers (for example KNM-ER 1805, KNM-ER 1813, and others) have moderate to extensive inflation of the supramastoid and nuchal crests. This feature, along with others suggests considerable variation in the hypodigm of *H. habilis*, and that more than one species may be represented in that hypodigm.

(ALD86-1) from a 4-m thick, bedded, grey, pumiceous tuff^{9,21} exposed at the hominid site. This unit occurs about 2.5 m below the specific horizon in the sequence that, on the basis of location and matching matrix, Martyn and Tobias indicated as the provenance of the hominid fossil⁹. Replicate ⁴⁰Ar/³⁹Ar single-crystal dates from the two samples gave identical weighted mean ages of 2.45 ± 0.02 Myr, for an overall age of 2.45 ± 0.01 Myr (Table 2). An isochron of these data (Fig. 1b) gives 2.43 ± 0.02 Myr, within error of the weighted mean age. Preliminary age results from tuffaceous units higher in the sequence (A.H., S.W. and A.D., manuscript in preparation) are consistent with these results, and support an age of about 2.4 Myr for the hominid. The isochron age is preferred because of the more flexible incorporation of the atmospheric correction.

Our work demonstrates that this part of the Chemeron Formation, contiguous with the type section^{9,21,22}, and laterally equivalent to it, is consistently younger than fossiliferous Chemeron outcrops elsewhere, such as Tabarin, which have also produced hominids^{23–26}. Our extensive fossil fauna from this section of the Formation (Table 3) accords with radioisotopic dates for the sequence ranging in age from 3.2 Myr to 1.6 Myr (A.H. *et al.*, manuscript in preparation).

Previously the oldest specimens reliably attributed to *Homo* were <1.9 Myr from East Turkana, Kenya^{27,28}, although some isolated teeth from the Shungura Formation, Ethiopia, possibly

as old as 2.4 Myr, are tentatively attributed to the genus by Howell *et al.*⁸ KNM-BC1 is therefore now important in extending the origin of the *Homo* clade more surely by about half a million years. In addition, the existence of *Homo* at 2.4 Myr is interesting because it is broadly contemporaneous with the earliest known stone tools and these occurrences fall close to a global cooling event.

The beginning of the archaeological record, in the form of the first lithic artefacts, is generally assumed to coincide with the origin of the genus *Homo*^{4,6,29,30}, but until now there has been a disjunction of up to 0.5 Myr between these two phenomena. The oldest known stone tools, assigned to the Oldowan Industrial Complex, come from Hadar, Ethiopia, palaeomagnetically dated at >2.4 Myr, and <2.6 Myr based on correlative radiometric determinations^{6,7}. Possibly slightly younger occurrences from members E and F of the Shungura formation are dated between 2.33 ± 0.03 Myr and 2.4 ± 0.05 Myr^{8,27,31,32}. Members of the *A. boisei* clade are also known as old as 2.50 Myr at West Turkana (KNM-WT 17000), and maybe 2.6 Myr in the Shungura formation (O.18-1967-18)^{18,27}. These therefore remain plausible candidates for artefact manufacture and use (see also refs 33, 34). The identification of *Homo* nearly contemporaneous with the earliest known artefacts makes more tenable the idea that our genus is also or exclusively responsible for the origins of lithic culture.

TABLE 2 $^{40}\text{Ar}/^{39}\text{Ar}$ data

Laboratory identification number	$^{40}\text{Ar}/^{39}\text{Ar}$	$^{37}\text{Ar}/^{39}\text{Ar}$	$^{36}\text{Ar}/^{39}\text{Ar}$	$^{40}\text{Ar}^*/^{39}\text{Ar}$	% $^{40}\text{Ar}^*$	Age (Myr $\pm 1\sigma$)
ALD86-1:						
395-01	0.5207	0.0105	0.00011	0.4812	92.4	2.43 $\pm$ 0.05
395-03	0.5566	0.0518	0.00021	0.4894	87.9	2.47 $\pm$ 0.15
395-04	0.5272	0.0300	0.00008	0.4980	94.5	2.52 $\pm$ 0.10
395-05	0.5132	0.0138	0.00013	0.4682	91.2	2.37 $\pm$ 0.04
395-06	0.5151	0.0104	0.00003	0.4992	96.9	2.52 $\pm$ 0.04
395-07	0.5611	0.0120	0.00020	0.4938	88.0	2.50 $\pm$ 0.04
395-08	0.5265	0.0133	0.00012	0.4844	92.0	2.45 $\pm$ 0.02
395-09	0.5136	0.0108	0.00008	0.4833	94.1	2.44 $\pm$ 0.02
					Weighted mean age	2.45 $\pm$ 0.02
Rejected						
395-02	1.5405	0.0250	0.003612	0.4664	30.3	2.36 $\pm$ 0.12
RD-6:						
396-01	0.5328	0.0215	0.00018	0.4738	88.9	2.40 $\pm$ 0.04
396-02	0.5355	0.0192	0.00015	0.4835	90.3	2.44 $\pm$ 0.03
396-03	0.5274	0.0201	0.00020	0.4608	87.4	2.33 $\pm$ 0.06
396-04	0.6656	0.0159	0.00056	0.4933	74.1	2.49 $\pm$ 0.03
396-05	0.5792	0.0202	0.00028	0.4886	84.4	2.47 $\pm$ 0.03
					Weighted mean age	2.45 $\pm$ 0.02
					Weighted mean age, ALD86-1 and RD-6	2.45 $\pm$ 0.01

$^{40}\text{Ar}/^{39}\text{Ar}$ analytical data on anorthoclase phenocrysts from a tuff exposed in the Kapthurin River valley, Tugen Hills, Kenya. Samples were irradiated for 10 MWh in the central thimble of the TRIGA reactor of the University of California, Berkeley. The flux monitor was hornblende mineral standard MMhb-1 with a K-Ar age of 520.4 ± 1.7 Myr³⁷. $^{40}\text{Ar}/^{39}\text{Ar}$ dating was done with a continuous Ar-ion laser under ultra-high vacuum. Argon isotope abundance ratios were measured with a MAP Series 215 noble-gas mass spectrometer. Corrections for inferring isotopes produced as a byproduct of irradiation were typically $<0.2\%$ of the calculated age for $^{36}\text{Ar}_{\text{Ca}}$, and $<2\%$ for $^{40}\text{Ar}_{\text{K}}$. The atmospheric ^{40}Ar component was, with one exception, less than 25%, and usually less than 10% of the total measured ^{40}Ar .

$\lambda = 5.543 \times 10^{-10} \text{ yr}^{-1}$; $J = 2.805 \times 10^{-3} \pm 9 \times 10^{-6}$; $(^{40}\text{Ar}/^{39}\text{Ar})_{\text{K}} = 0.0086$, $(^{36}\text{Ar}/^{37}\text{Ar})_{\text{Ca}} = 2.59 \times 10^{-4}$, $(^{39}\text{Ar}/^{37}\text{Ar})_{\text{Ca}} = 9.0 \times 10^{-4}$; errors for the weighted mean are according to Samson and Alexander³⁷.

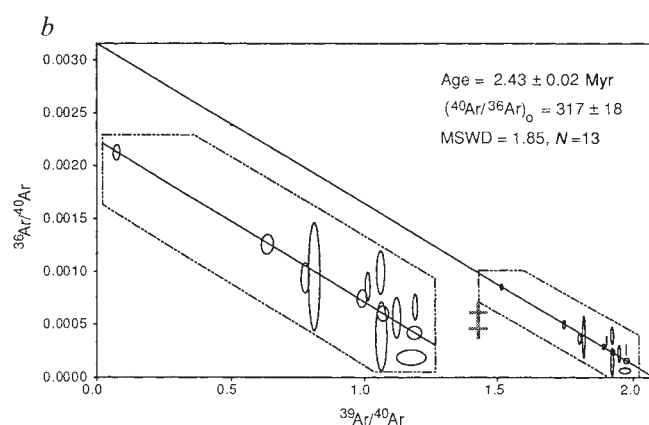
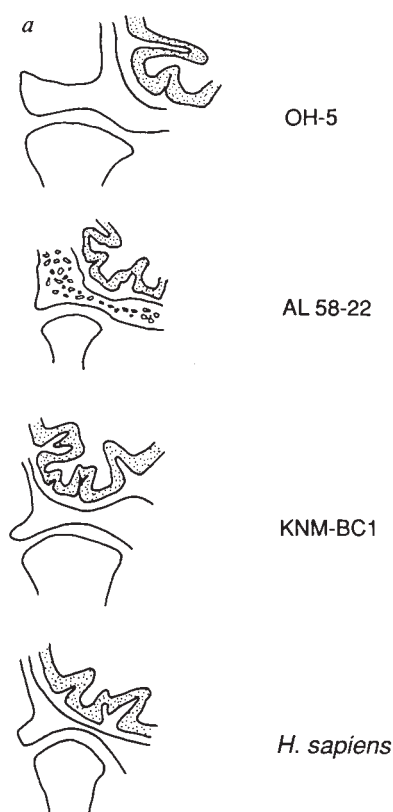


FIG. 1 *a*, Schematic reconstructions of hominid temporal bones of *A. boisei* (OH5), *A. afarensis* (AL 85-22), KNM-BC1 and *H. sapiens*. Coronal sections through the floor of the mandibular fossa. Positions of hypothetical temporal lobes of the cerebrum and mandibular condyles are indicated (stippled). *b*, $^{36}\text{Ar}/^{40}\text{Ar}$ versus $^{39}\text{Ar}/^{40}\text{Ar}$ isotope correlation diagram of samples ALD86-1 and RD-6, from the 'Lower Tuffs' at site K002/JM85. Error ellipses are correlated 1σ analytical uncertainties. The diagonal line is the weighted least-squares solution to the data using a York fitting approach, incorporating errors on both axes and the correlation of x and y errors³⁹. The arrow points toward an enlargement of the data set. The atmospheric $^{40}\text{Ar}/^{39}\text{Ar}$ composition of 317 ± 18 indicated by the isochron is within 2σ of the true air $^{40}\text{Ar}/^{39}\text{Ar}$ of 295.5.

TABLE 3 Chemeron Formation faunal list

Pisces		
Siluriformes	Clariidae	
	Cichlidae	
Reptilia		
Chelonia	Testudinidae	<i>Geochelone</i> sp.
	Trionychidae	
Crocodylia	Crocodylidae	<i>Crocodylus</i> sp.
Mammalia		
Primates	Cercopithecidae	<i>Paracolobus chemeroni</i> <i>Papio baringensis</i> <i>Theropithecus brumpti</i> <i>Homo</i> sp.
	Hominidae	
Carnivora	Hyenidae	<i>Crocota</i> cf. <i>C. crocuta</i>
	cf. Viverridae	
Tubulidentata	Orycteropodidae	<i>Orycteropus</i> sp.
Proboscidea	Deinotheriidae	<i>Deinotherium bozasi</i>
	Elephantidae	<i>Elephas recki shungurensis</i> <i>Loxodonta</i> cf. <i>L. exoptata</i>
Perissodactyla	Rhinocerotidae	<i>Ceratotherium</i> cf. <i>C. simum</i>
	Equidae	<i>Hipparion</i> cf. <i>H. ethiopicum</i>
	Chalicotheriidae	cf. <i>Ancylotherium</i>
Artiodactyla	Suidae	<i>Kolpochoerus afarensis</i> <i>Metridiochoerus andrewsi</i> <i>Notochoerus scotti</i> <i>Hexaprotodon imagunculus</i> <i>Hippopotamus</i> cf. <i>H. kaisensis</i>
	Hippopotamidae	<i>Giraffa</i> sp.
	Giraffidae	<i>Sivatherium maurusium</i>
	Bovidae	
	Tragelaphini	<i>Tragelaphus</i> sp. (large)
	Reduncini	cf. <i>Kobus sigmoidalis</i>
	Alcelaphini	
	Bovini	cf. <i>Pelorovis</i>
	Caprini	
	Antilopini	cf. <i>Antidorcas recki</i>

Faunal list for the Chemeron Formation (*sensu stricto*), based on the many sites now known from the type section and other correlative sediments, which span 3.2 Myr to 1.6 Myr. The hominid site (BPRP number K002; JM85) is in lacustrine sediments with abundant fish fossils. But we have also collected many mammals from this site, some *in situ*, showing the same preservation as the hominid. From the original collections Gentry³⁸ documented five bovid fossils from the site, all belonging to Alcelaphini. We now have other bovid specimens, and other taxa from K002/JM85 include Hippopotamidae, Suidae and Proboscidea.

Although we urge caution in placing great significance on the simple coincidence of different classes of data preserved at different levels of resolution in the fossil record, it is intriguing to correlate the origin of our genus with global climatic shifts, possibly extraterrestrially triggered. Some recent environmentally deterministic hypotheses of speciation use origins of taxa in Hominidae as examples of a causal relation between speciation and climatic change. Vrba^{3,4} suggests global cooling at 2.4 Myr as a factor in the origin of *Homo*, a datum she establishes unsatisfactorily on the oldest lithic artefacts. The existence of *Homo* at 2.4 Myr makes these notions more plausible, although problems still exist³⁵.

The proposed diagnostic features for *Homo* described here will be useful in elucidating the relationships of other enigmatic hominid specimens, and will help in defining a currently poorly defined taxon. *Homo* at 2.4 Myr also has phylogenetic implications. □

Received 16 October; accepted 21 November 1991.

- Thunell, R. C. & Williams, D. F. *Utrecht Micropaleont. Bull.* **30**, 111-127 (1983).
- Shackleton, N. J. *et al. Nature* **307**, 620-623 (1984).
- Vrba, E. S. *Afr. J. Sci.* **81**, 229-236 (1985).
- Vrba, E. in *Evolutionary History of the 'Robust' Australopithecines* (ed. Grine, F.), 405-426 (Aldine de Gruyter, New York, 1988).
- Prentice, M. L. & Denton, G. H. in *Evolutionary History of the 'Robust' Australopithecines* (ed. Grine, F.) 383-403 (Aldine de Gruyter, New York, 1988).

- Roche, H. & Tiercelin, J.-J. in *Proc. 8th Panafrican Congress of Prehistory and Quaternary Studies* (eds Leakey, R. E. & Ogot, B. A.) 194-199 (KNM, Nairobi, 1980).
- Harris, J. *Afr. archaeol. Rev.* **1**, 3-31 (1983).
- Howell, F. C. *et al. J. hum. Evol.* **16**, 665-700 (1987).
- Martyn, J. & Tobias, P. V. *Nature* **215**, 476-480 (1967).
- Leakey, L. S. B. & Leakey, M. D. Excavation report, Bishop Archives, National Museums of Kenya.
- Leakey, R. *Symp. zool. Soc. Lond.* **33**, 53-69 (1973).
- Oakley, K., Campbell, B. & Molleson, T. (eds) *Catalogue of Fossil Hominids: Africa* 223 (British Museum, London, 1977).
- Kimbel, W. H. & Rak, Y. *Am. J. phys. Anthropol.* **66**, 31-54 (1985).
- Kimbel, W. H. thesis, Kent State University (1986).
- Dean, C. in *Evolutionary History of the 'Robust' Australopithecines* (ed. Grine, F.) 107-112 (Aldine de Gruyter, New York, 1988).
- Sherwood, R. J. & Ward, S. *Am. J. phys. Anthropol.* **78**, 301 (1989).
- Picq, P. thesis, Univ. Paris (1985).
- Kimbel, W. H. & White, T. D. in *Evolutionary History of the 'Robust' Australopithecines* (ed. Grine, F.) 175-192 (Aldine de Gruyter, New York, 1988).
- Kimbel, W. H., White, T. D. & Johanson, D. C. in *Evolutionary History of the 'Robust' Australopithecines* (ed. Grine, F.) 259-268 (Aldine de Gruyter, New York, 1988).
- Leakey, R. E. F. & Walker, A. C. *Am. J. phys. Anthropol.* **76**, 1-24 (1988).
- Martyn, J. thesis, Univ. London (1969).
- McCall, G. J. H. *et al. in Background to Evolution in Africa* (eds Bishop, W. W. & Clark, J. D.) 191-220 (University of Chicago Press, 1967).
- Pickford, M. *et al. Am. J. phys. Anthropol.* **60**, 337-346 (1983).
- Hill, A. *Nature* **315**, 222-224 (1985).
- Ward, S. & Hill, A. *Am. J. phys. Anthropol.* **72**, 21-37 (1987).
- Hill, A. & Ward, S. *Yb phys. Anthropol.* **31**, 49-83 (1988).
- Feibel, C. G. *et al. Am. J. phys. Anthropol.* **78**, 595-622 (1989).
- Leakey, M. G. & Leakey, R. E. F. (eds) *Koobi Fora Research Project Vol. 1*, 191 (Oxford University Press, 1978).
- Isaac, G. *Scient. Am.* **238**, 90-108 (1978).
- Harris, J. in *The Cambridge Encyclopedia of Archaeology* (ed. Sherratt, A.) 62-70 (Cambridge University Press, Cambridge, 1980).
- Merrick, H. *et al. Nature* **242**, 572-575 (1973).
- Merrick, H. in *Human Origins: Louis Leakey and the East African Evidence* (eds Isaac, G. & McCown, E.) 461-481 (Benjamin, Menlo Park, 1976).
- Susman, R. L. *Science* **240**, 781-784 (1988).
- Susman, R. L. in *Evolutionary History of the 'Robust' Australopithecines* (ed. Grine, F.) 149-172 (Aldine de Gruyter, New York, 1988).
- Hill, A. *J. hum. Evol.* **16**, 583-596, (1987).
- Walker, A. & Leakey, R. E. F. in *Evolutionary History of the 'Robust' Australopithecines* (ed. Grine, F.) 247-258 (Aldine de Gruyter, New York, 1988).
- Samson, S. D. & Alexander, E. C. *Chem. geol. Isotope Geosci.* **66**, 27-34 (1987).
- Gentry, A. in *Geological Background to Fossil Man* (ed. Bishop, W. W.) 293-308, (Scottish Academic, Geol. Soc. Lond., 1978).
- York, D. *Earth Planet. Sci. Lett.* **5**, 320-324 (1969).

ACKNOWLEDGEMENTS. We thank J. Kimengich, J. Martyn, Mary Leakey, R. Leakey and A. Walker for their help and advice regarding this specimen. E. Mbuu for assistance in Nairobi, L. Budinoff for preparing Fig. 1 and other members of the project for their help. The National Museums of Kenya provides much logistical support, and we thank the Government of the Republic of Kenya for research permission. The Baringo Paleontological Research Project, based at Yale University, operating jointly with the National Museums of Kenya, is funded by the NSF, the Louise Brown Foundation, the Ingalls Foundation, the LSB Leakey Foundation, Mr C. Stephenson and the Yale University Social Science Faculty Research Fund. S.W. also acknowledges support from the NSF.

Potentialiation of NMDA receptor currents by arachidonic acid

Barbara Miller, Monique Sarantis, Stephen F. Traynelis* & David Attwell

Department of Physiology, and * Department of Pharmacology, University College London, Gower Street, London WC1E 6BT, UK

ARACHIDONIC acid is released by phospholipase A₂ when activation of *N*-methyl-D-aspartate (NMDA) receptors by neurotransmitter glutamate raises the calcium concentration in neurons^{1,2}, for example during the initiation of long-term potentiation and during brain anoxia^{3,4}. Here we investigate the effect of arachidonic acid on glutamate-gated ion channels by whole-cell clamping isolated cerebellar granule cells. Arachidonic acid potentiates, and makes more transient, the current through NMDA receptor channels, and slightly reduces the current through non-NMDA receptor channels. Potentiation of the NMDA receptor current results from an increase in channel open probability, with no change in open channel current. We observe potentiation even with saturating levels of agonist at the glutamate- and glycine-binding sites on these channels; it does not result from conversion of arachidonic acid to lipoxygenase or cyclooxygenase derivatives, or from activation of protein kinase C. Arachidonic acid may act by binding to a site on the NMDA receptor, or by modifying the receptor's lipid environment. Our results suggest that arachidonic acid released by activation of NMDA (or other) receptors will potentiate NMDA

receptor currents, and thus amplify increases in intracellular calcium concentration caused by glutamate. This may explain why inhibition of phospholipase A₂ blocks the induction of long-term potentiation⁵.

Cerebellar granule cells express both NMDA- and non-NMDA-type glutamate-gated channels. Currents through these channels were monitored while repeatedly superfusing solution containing 100 μ M kainate (to activate non-NMDA receptors) or 15 μ M aspartate plus 5 μ M glycine (to activate NMDA receptors). Application of 5 μ M arachidonic acid either had no effect (Fig. 1a) or depressed the current through non-NMDA receptor channels. In seven cells, the kainate-evoked current in arachidonic acid was reduced by 20 ± 7 (s.e.)%. Similarly, the initial peak of the current evoked by (S)- α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) iontophoresis was reduced by $28 \pm 9\%$ (four cells; data not shown). Non-NMDA receptor currents were not studied further. In contrast, the peak current through NMDA receptor channels was more than doubled by 5 μ M arachidonic acid (Fig. 1a).

The potentiation of aspartate-evoked NMDA receptor currents in Fig. 1a cannot be due to arachidonic acid increasing the local aspartate concentration by inhibiting aspartate uptake⁶, because the same potentiation was seen (1) for currents evoked by NMDA, which is not taken up⁷ (see Fig. 2 legend); (2) for cells lifted off the bottom of the culture dish and so removed from the effect of uptake in nearby cells (Fig. 1f-h); and (3) for currents evoked by a saturating dose of aspartate (see later).

To examine the effect of arachidonic acid with higher temporal resolution, aspartate or NMDA were applied by iontophoresis (Fig. 1b). Arachidonic acid (10 μ M) increased the initial peak of the aspartate- or NMDA-induced current (quantified in Fig. 2c), decreased its time-to-peak (by $8.2 \pm 3.5\%$ from a control value of 103 ± 9 ms in 14 cells), and made the current more transient (Fig. 1c-e). The ratio of the peak NMDA-evoked current to the current at the end of the 540-ms iontophoresis pulse had a control value of 2.63 ± 0.28 in 14 cells, and was increased to 4.72 ± 0.77 by 10 μ M arachidonic acid. Arachidonic acid potentiated the peak current to the same extent, independently of whether the current waveform was initially relatively non-transient ($I_{\text{peak}}/I_{540 \text{ ms}} \approx 1$, as in Fig. 1c) or more transient ($I_{\text{peak}}/I_{540 \text{ ms}}$ up to 6). The increase of transience produced by arachidonic acid contrasts with the effect of glycine, which potentiates NMDA receptor currents but (in some⁸ but not all⁹ situations) makes them less transient.

In some cells, aspartate or NMDA opened only a few channels. The low capacitance and high resistance of granule cells provides high enough resolution in whole-cell mode of single-channel currents to allow us to test whether arachidonic acid increases NMDA receptor currents by increasing the current through open channels, or by increasing the channel open probability. Most channel openings in normal solution were to states with a conductance of ~ 55 pS: in Fig. 1f and h, iontophored NMDA results in a maximum of three or four such channels being open simultaneously. Applying 5 μ M arachidonic acid leads to more channels opening at the peak current (up to seven in Fig. 1g), with essentially no change in open channel current. For the cell used for Fig. 1f-h, the mean current change at -68 mV in 200 NMDA-evoked single channel transitions in the control solution was 3.66 ± 0.03 (s.e.) pA, whereas in arachidonic acid the corresponding value was 3.54 ± 0.03 pA. Similar results were obtained from six other cells. Thus, arachidonic acid potentiates NMDA receptor currents by increasing the number of channels open, either by increasing the opening rate or mean open time of all channels, or by recruiting previously inactive channels. The increase in current transience and decreased time to peak current in arachidonic acid are consistent with an increase in rate of one of the steps leading to channel opening.

Arachidonic acid had no significant effect on the NMDA receptor current reversal potential (Fig. 2a). Currents above and below the reversal potential were potentiated to a similar degree

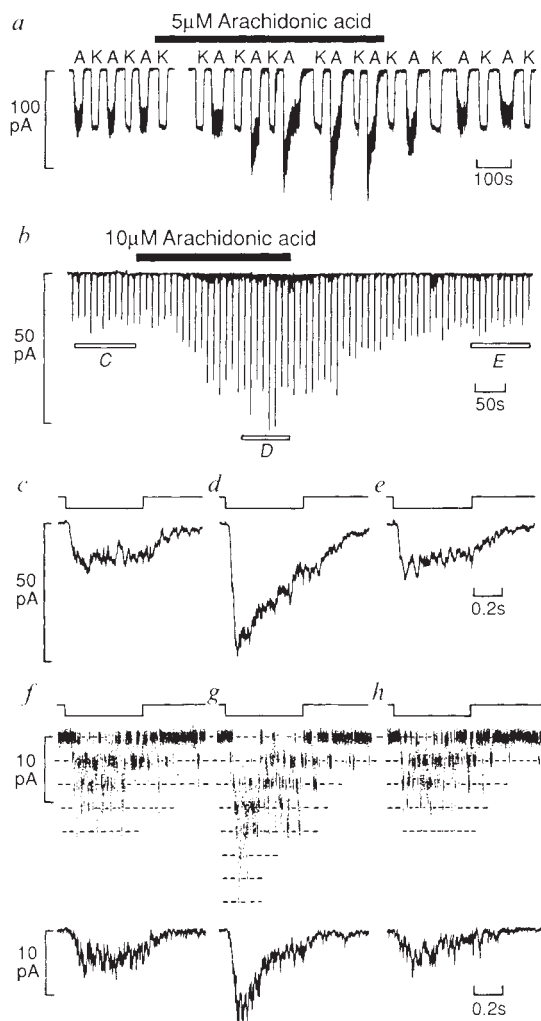
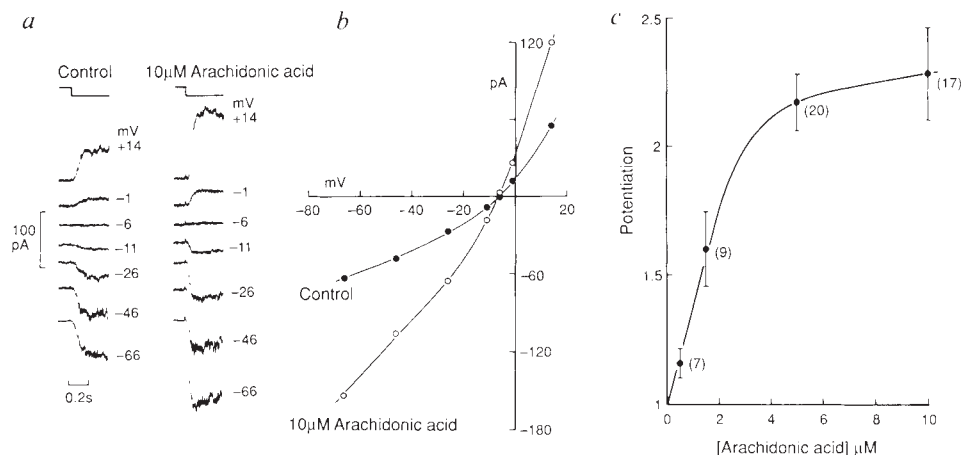


FIG. 1 Arachidonic acid potentiates NMDA receptor currents. *a*, Membrane current at -66 mV of a cerebellar granule cell during application of 15 μ M aspartate plus 5 μ M glycine (A) or 100 μ M kainate (K) to activate NMDA and non-NMDA receptors selectively^{10,24}. *b*, Membrane current at -66 mV of a cell to which L-aspartate was repeatedly iontophored: each downward deflection is the inward NMDA receptor current. *c*, *d* and *e*, Averaged responses for the periods shown as *C*, *D* and *E* in *b*, before, during and after application of 10 μ M arachidonic acid. Top traces show iontophoretic current. *f*, *g* and *h*, NMDA-evoked currents at -68 mV from an experiment like that in *b* but on a cell pulled off the bottom of the bath and showing only a few NMDA receptor channels. Top traces: single responses to iontophoresis before (*f*), during (*g*) and after (*h*) 5 μ M arachidonic acid. Dashed lines are separated by 3.6 pA. Bottom traces: averages of 6 responses before, during and after arachidonic acid show a change in current waveform similar to that seen in cells with more responding channels (*c-e*).

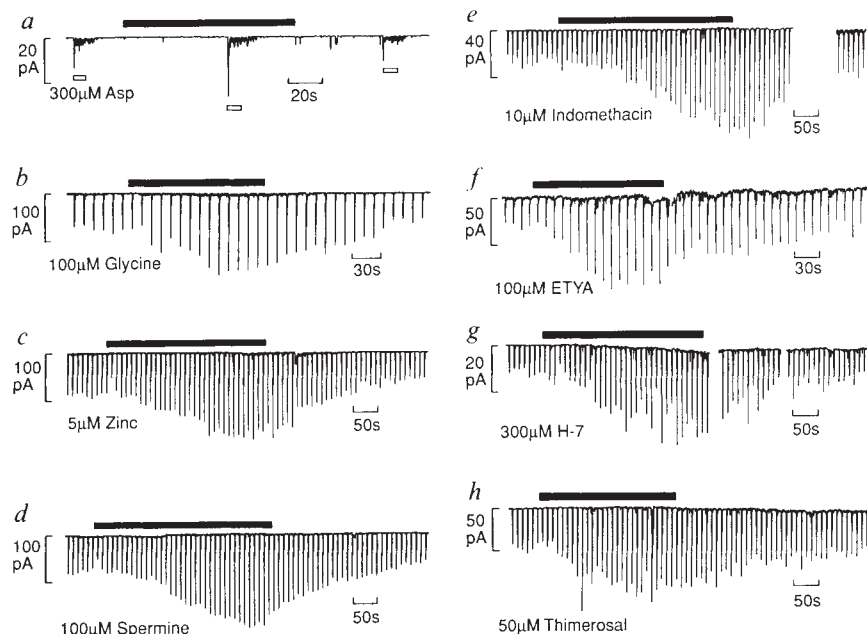
METHODS. Granule cells were obtained from explant cerebellar cultures from 7-day-old rats²⁵. External solution contained 5 μ M glycine, except where stated otherwise, and (in mM): NaCl, 143; KCl, 2.8; CaCl₂, 1; HEPES, 10; glucose, 15; pH adjusted to 7.3 with NaOH. Pipette solution contained (in mM): CsF, 110; CsCl, 30; CaCl₂, 0.5; K₂EGTA, 5; HEPES, 5; pH adjusted to 7.2 with KOH. Calcium buffering by EGTA and F⁻ in this solution may prevent release of endogenous arachidonic acid when NMDA receptors are activated^{1,2}. Pipette junction potential of -6 mV has been added to all voltages. Pipette series resistance in whole-cell mode was typically 50 M Ω . Arachidonic acid was made up from a stock in organic solvent (ethanol) as described in ref. 16, or in water⁶, which gave the same results. To estimate voltage-clamp quality, cells were assumed to comprise a 7- μ m diameter soma attached to a 50- μ m-long, 1- μ m diameter neurite with a sealed end and internal resistivity 2 Ω m. For a typical input resistance of 2 G Ω , assumed to be uniformly distributed over the cell surface area, this leads to a dendrite space constant of 280 μ m, and $<2\%$ voltage non-uniformity along the neurite for voltage steps imposed at the soma. Arachidonic acid potentiated NMDA receptor currents even in cells lacking neurites. Temperature, 20–25 $^{\circ}$ C.

FIG. 2 Voltage- and dose-dependence of the potentiation by arachidonic acid. **a**, Specimen responses to iontophoresed NMDA at different membrane potentials (shown by each trace) before and during application of arachidonic acid. Reversal potential was -4.0 ± 0.9 mV ($n=12$) in the control solution and was shifted by -1.0 ± 0.4 mV ($n=5$) in $5\text{--}10$ μM arachidonic acid. **b**, Voltage-dependence of peak NMDA-evoked currents (from the cell in **a**) before and in arachidonic acid. In 6 cells, the potentiation of the peak NMDA-evoked current produced by 5 or 10 μM arachidonic acid at +14 mV was 1.2 ± 0.2 times the potentiation produced at -66 mV. **c**, Potentiation of peak NMDA receptor currents at -66 mV by different doses of arachidonic acid: currents were evoked by iontophoresis of NMDA with 5 μM external glycine, except where stated below. Control values of current (and other parameters reported) were measured by averaging values before and after arachidonic acid. The potentiation produced by 5 μM arachidonic acid with 5 μM glycine in the external solution (peak current increased by a factor of 2.19 ± 0.15 , $n=11$ cells) was not significantly different from that produced with 100 μM glycine present (2.15 ± 0.18 , $n=5$), or from that produced with 5 μM glycine and 100 μM ascorbic acid to guard against arachidonic acid oxidation (2.14 ± 0.33 , $n=4$), and these data were pooled for the 5 μM point



on this dose-response curve. For 10 μM arachidonic acid, the potentiation produced with NMDA as the iontophoresed glutamate analogue (2.09 ± 0.21 , $n=9$) was not significantly different from that for aspartate-evoked currents (2.59 ± 0.34 , $n=7$) and these data were pooled, together with data from one cell studied using NMDA and 100 μM glycine, to obtain the 10 μM point on the curve. The increase of current transience (assessed as the ratio of the peak current to the current at the end of the iontophoresis pulse) had a dependence on arachidonic acid dose similar to that for the potentiation of the peak current shown here.

FIG. 3 Pharmacological analysis of the potentiation by 5 μM arachidonic acid (heavy overbars). Currents were evoked at -66 mV, by iontophoresis of NMDA (except for **a**), and with 5 μM glycine present (except for **b**). In all of the manipulations shown, the potentiation of the peak current produced by arachidonic acid was not significantly reduced from the control value (2.19 ± 0.15 (s.e.), $n=11$; see Fig. 2c legend) seen for NMDA iontophoresis with 5 μM external glycine. **a**, Potentiation of the current evoked by 300 μM L-aspartate (open bars). Peak current was potentiated by a factor 2.16 ± 0.14 in 4 cells. **b**, Potentiation with 100 μM glycine present. Peak current was potentiated by 2.15 ± 0.18 ($n=5$). **c**, Potentiation in the presence of 5 μM zinc, which roughly halves the NMDA-evoked current by decreasing the open probability of the NMDA-gated channel¹². Potentiation of peak current was 2.57 ± 0.39 ($n=3$). **d**, Potentiation in the presence of 100 μM spermine, which roughly doubles the NMDA-evoked current¹³. Potentiation of peak current by arachidonic acid was 2.41 ± 0.42 ($n=6$). **e**, Potentiation in the presence of 10 μM indomethacin (applied 5 min before arachidonic acid). Potentiation of peak current was 2.58 ± 0.29 ($n=5$). IC_{50} value¹⁴ for indomethacin blocking cyclooxygenase is <0.1 μM . **f**, Potentiation with 100 μM ETYA included in the pipette solution. Arachidonic acid was applied 23 min (in other cells 10–33 min) after going to whole-cell mode. Series resistance was 56 M Ω , leading to an exchange time constant of 31 s for a molecule of diffusion coefficient 5×10^{-10} $\text{m}^2 \text{s}^{-1}$ (the value for sucrose²⁶ which has a similar molecular weight to ETYA and H-7) entering a cell of volume 219 μm^3 from a pipette filled with solution of resistivity 0.8 Ωm . Diffusion time down a 50- μm -long neurite is about 5 s. Potentiation of peak current was 2.71 ± 0.41 ($n=5$). IC_{50} values for ETYA blocking 12- and 15-lipoxygenase, 5-lipoxygenase, cyclooxygenase and epoxigenase are 0.3 μM , 10 μM , <8 μM , and 40 μM respectively^{14,15}. **g**, Potentiation with 300 μM H-7 (1-(5-isoquinolinesulphonyl)-2-methyl-piperazine) in the pipette



solution. Arachidonic acid was applied 5 min after going to whole-cell mode. The series resistance of 60 M Ω predicts an exchange time constant of 33 s for H-7 entering the cell from the pipette. Potentiation of peak current was 2.44 ± 0.51 ($n=5$) (or 2.77 ± 0.22 ($n=4$) with 100 nM pipette staurosporine). IC_{50} s for H-7 and staurosporine blocking protein kinase C are 6 μM and 3 nM (ref. 17). **h**, Potentiation in the presence of 50 μM thimerosal (merthiolate), which inhibits¹⁹ incorporation of arachidonic acid into phospholipids and triglycerides by $>90\%$. Arachidonic acid was added after 2 h (>35 min in other cells) in thimerosal. Potentiation of peak current was 2.13 ± 0.11 ($n=5$).

(Fig. 2b). Potentiation of NMDA receptor currents was seen at low micromolar levels of arachidonic acid and was nearly saturated, with a more than twofold increase in peak current, at 10 μ M arachidonic acid (Fig. 2c). Potentiation of NMDA receptor currents was also seen in putative pyramidal cells acutely isolated by trypsin dissociation of adult rat hippocampal slices: 5 μ M arachidonic acid increased the peak NMDA-evoked current (in 5 μ M glycine) by a factor of 2.12 ± 0.02 (s.e.) in three cells, similar to the increase seen in cerebellar cells.

The potentiating effect of arachidonic acid cannot be due simply to arachidonic acid increasing the affinity of the NMDA receptor for agonists acting at its glutamate or glycine binding sites: similar potentiation was seen for currents evoked by 300 μ M aspartate (Fig. 3a) or in the presence of 100 μ M glycine (Fig. 3b), which saturate these sites^{8,10}. The nominal absence of Mg^{2+} in the external solution and the lack of strong voltage-dependence to the potentiation (Fig. 2b) suggest that arachidonic acid is not acting by modulating divalent cation binding to the magnesium-binding site on the NMDA receptor¹¹. The presence of 5 μ M zinc (Fig. 3c) or 100 μ M spermine (Fig. 3d) did not significantly affect the potentiation produced by arachidonic acid, consistent with arachidonic acid acting independently of the zinc¹² and polyamine¹³ binding sites on the NMDA receptor.

We tested whether a second messenger system might generate the potentiation produced by arachidonic acid. In the presence of indomethacin, which is a cyclooxygenase blocker¹⁴, or 5,8,11,14-eicosatetraenoic acid (ETYA), which blocks lipoxygenase, cyclooxygenase and epoxigenase^{14,15}, arachidonic acid still potentiated the NMDA receptor current (Fig. 3e and f), implying that it does not act by being converted to prostaglandins or metabolites like hydroperoxyeicosatetraenoic acids (HPETEs)¹⁶, hydroxyeicosatetraenoic acids (HETEs) or epoxyeicosatrienoic acids (EETs). The protein kinase inhibitors¹⁷ H-7 (Fig. 3g) and staurosporine (not shown) did not prevent the potentiation by arachidonic acid, showing that activation¹⁸ of protein kinase C is not involved. The acyl-transferase blocker thimerosal¹⁹ also had no effect on the potentiation (Fig. 3h), implying that arachidonic acid does not act by being incorporated into membrane phospholipids and decreasing the concentration of lysophospholipids. Arachidonic acid could act by binding directly to a novel site on the NMDA receptor or by altering the receptor's lipid environment.

These results suggest that NMDA receptors could respond to arachidonic acid released when glutamate activates NMDA receptors in the same or in neighbouring neurons. As a result, synaptically released glutamate will be able to open more NMDA receptor channels, leading to more arachidonic acid release and an amplification of changes of calcium concentration produced in neurons by glutamate. Arachidonic acid released by basal, or non-glutamate receptor-induced²⁰, activity of phospholipase A_2 could also potentiate synaptically evoked current flow and calcium entry through NMDA receptor channels. Interestingly, inhibition of phospholipase A_2 to prevent arachidonic acid release prevents the induction of long-term potentiation (LTP) at hippocampal area CA1 synapses, where LTP is initiated by a calcium influx through NMDA receptor channels (but does not prevent it at area CA3 mossy fibre synapses where LTP is not initiated by NMDA receptors)^{5,21,22}. Inhibition of phospholipase A_2 could curtail LTP in area CA1 by decreasing calcium entry through NMDA receptor channels, as the duration of LTP depends critically on the amount of calcium entry²³. □

7. Sarantis, M. & Attwell, D. *Brain Res.* **516**, 322–325 (1990).
8. Mayer, M. L., Vyklicky, L. Jr & Clements, J. *Nature* **338**, 425–427 (1989).
9. Sather, W., Johnson, J. W., Henderson, G. & Ascher, P. *Neuron* **4**, 725–731 (1990).
10. Patneau, D. K. & Mayer, M. L. *J. Neurosci.* **10**, 2385–2399 (1990).
11. Nowak, L., Bregestowski, P., Ascher, P., Herbet, A. & Prochiantz, A. *Nature* **307**, 462–465 (1984).
12. Christine, C. W. & Choi, D. W. *J. Neurosci.* **10**, 108–116 (1990).
13. Sprosen, T. S. & Woodruff, G. N. *Eur. J. Pharmacol.* **179**, 477–478 (1990).
14. Salari, H., Braquet, P. & Borgeat, P. *Prostaglandins, Leukotrienes & Medicine* **13**, 53–60 (1984).
15. Capdevila, J. et al. *Arch. Biochem. Biophys.* **261**, 257–263 (1988).
16. Piomelli, D. et al. *Nature* **328**, 38–43 (1987).
17. Rüegg, U. T. & Burgess, G. M. *Trends Pharmacol. Sci.* **10**, 218–220 (1989).
18. Sekiguchi, K., Tsukuda, M., Ase, K., Kikkawa, U. & Nishizuka, Y. *J. Biochem. (Tokyo)* **103**, 759–765 (1988).
19. Goppelt-Streube, M., Koerner, C.-F., Hausmann, G., Gerns, D. & Resch, K. *Prostaglandins* **32**, 373–385 (1986).
20. Axelrod, J., Burch, R. M. & Jelsema, C. L. *Trends Neurosci.* **11**, 117–123 (1988).
21. Williams, J. H. & Bliss, T. V. P. *Neurosci. Lett.* **88**, 81–85 (1988).
22. Linden, D. J., Sheu, F. S., Murakami, K. & Routenberg, A. *J. Neurosci.* **7**, 3783–3792 (1987).
23. Malenka, R. C. *Neuron* **6**, 53–60 (1991).
24. Davies, J., Evans, R. H., Jones, A. W., Smith, D. A. S. & Watkins, J. C. *Comp. Biochem. Physiol.* **72C**, 211–224 (1982).
25. Cull-Candy, S. G., Howe, J. R. & Ogden, D. C. *J. Physiol.* **400**, 189–222 (1988).
26. *CRC Handbook of Chemistry & Physics* (1971–1972) 52nd edn (ed. Weast, R. C.) F-47 (Chemical Rubber Co., 1971).

ACKNOWLEDGEMENTS. Supported by the Wellcome Trust, the MRC, the NIH and the Wolfson Foundation. We thank J. Thomas for culturing the cerebellar cells, G. Christofi for hippocampal slices for dissociation, and S. Cull-Candy, A. Gibb, T. Sanders and D. Wyllie for advice.

Impaired type II glucocorticoid-receptor function in mice bearing antisense RNA transgene

Marie-Claude Pepin, François Pothier* & Nicholas Barden

Molecular Psychogenetics Laboratory, and Departments of Physiology and * Zootechnie, CHUL Research Center and Laval University, 2705 Boulevard Laurier, Sainte Foy, Québec G1V 4G2, Canada

GLUCOCORTICOIDS, in conjunction with their cognate receptors, exert negative-feedback effects on the hypothalamus–pituitary–adrenal axis, suppressing adrenal steroid secretions. Two types of corticosteroid receptor, distinguishable by their ability to bind corticosterone, have been identified as classical mineralocorticoid (type I)¹ and glucocorticoid (type II)² receptors by cloning their complementary DNAs. The type I receptor controls the basal circadian rhythm of corticosteroid secretion. Both receptor types are involved in negative feedback³, but the type II receptor may be more important for terminating the stress response as it is the only one to be increased in animals rendered more sensitive to corticosteroid negative-feedback effects⁴. Here we create a transgenic mouse with impaired corticosteroid-receptor function by partially knocking out gene expression with type II glucocorticoid receptor antisense RNA. We use this animal to study the glucocorticoid feedback effect on the hypothalamus–pituitary–adrenal axis.

We microinjected into mouse oocytes a type-II glucocorticoid receptor antisense RNA construct previously shown to decrease functional type II glucocorticoid receptor levels in cells in culture⁵. A fragment of 1,815 base pairs (bp) from the 3' non-coding region of rat type II glucocorticoid receptor cDNA, inverted downstream from a human neurofilament gene promoter element (pNFLAsGR), was used to minimize sequence homology and cross-inhibition of other members of the steroid hormone gene superfamily (Fig. 1). Of 13 mice screened for the presence of this transgene, five were positive for integration (lines 1.3, 5.2, 5.4, 5.7 and 5.9). All of these founders except founder 5.7, which had only one transgene copy and could have been a mosaic, transmitted the transgene to their offspring. Only mice derived from founder 5.2 had reduced reproductive capacity.

All the transgenic lines display a characteristic phenotype whether they are homozygotes or heterozygotes (Fig. 2a). Transgenic animals have greatly increased fat deposition and, at ~6 months of age, can be twice the weight of normal mice (Fig. 2b).

Received 21 August; accepted 22 November 1991.

1. Dumuis, A., Sebben, M., Haynes, L., Pin, J.-P. & Bockaert, J. *Nature* **336**, 68–70 (1988).
2. Lazarewicz, J. W., Wroblewski, J. T. & Costa, E. *J. Neurochem.* **55**, 1875–1881 (1990).
3. Lynch, M. A., Errington, M. L. & Bliss, T. V. P. *Neuroscience* **30**, 693–701 (1989).
4. Rehncrona, S., Westerberg, E., Akesson, B. & Siesjö, B. K. *J. Neurochem.* **38**, 84–93 (1982).
5. Okada, D., Yamagishi, S. & Sugiyama, H. *Neurosci. Lett.* **100**, 141–146 (1989).
6. Barbour, B., Sztakowski, M., Ingledew, N. & Attwell, D. *Nature* **342**, 918–920 (1989).

This fat accumulates in different adipose tissue regions, despite the fact that the transgenic mice eat 15% less than the normal mice (D. Richard *et al.*, unpublished results).

The mechanism of action of antisense RNA is unknown and so far there is but one example of antisense RNA inhibition of gene expression (myelin basic protein) in transgenic animals⁶. When injected directly into the cytoplasm of cells, antisense RNA must complex with the 5' region of the endogenous mRNA to block its translation⁷, but if expressed in the nucleus, antisense RNAs directed against either 5' or 3' ends of endogenous mRNA can be effective^{8,9}. Using amplification by reverse transcriptase with the polymerase chain reaction (PCR), we demonstrated the presence and uniqueness of antisense mRNA in tissue extracts from transgenic animals (Fig. 3a). Northern blot analysis of endogenous type II glucocorticoid receptor messenger RNA indicates a maximal 50–70% decrease in hypothalamus and frontal cortex of the three different transgenic lines analysed; a smaller 30–55% decrease is seen in liver (Fig. 3b). It is possible that hybrid RNA strands, once formed, are rapidly degraded by ribonucleases. This is supported by the decreased type II glucocorticoid receptor mRNA concentrations and by our in-

ability to detect by northern blotting the very low levels of antisense mRNA remaining in the cell using probes directed against various regions of the transgene. The decreased type II glucocorticoid receptor mRNA concentrations are reflected by decreased glucocorticoid binding activity in tissues where the transgene is expressed. In transgenic animals, glucocorticoid binding sites are diminished in frontal cortex, hypothalamus, pituitary gland and liver (Table 1). We had hoped that the human neurofilament-L gene promoter would restrict expression of the transgene to the brain, as was shown previously using the complete 21.5-kilobase (kb) gene clone¹⁰. Nucleotide sequences in

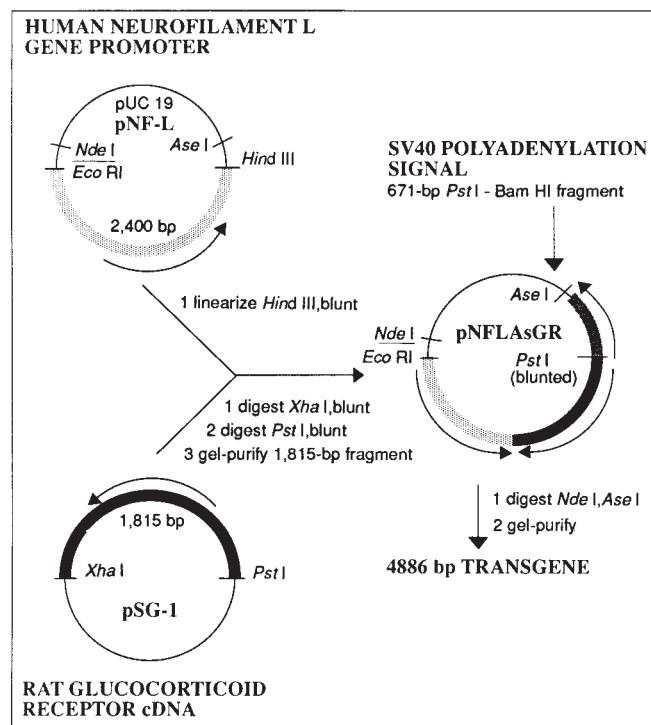


FIG. 1 Schematic representation of the structure and construction of the NFL-AsGR transgene used to direct the transcription of an antisense RNA with a sequence complementary to that of a section of rat type II glucocorticoid receptor mRNA. Plasmid pNFL-AsGR was constructed by cloning a 1,815-bp fragment of the 3' non-coding region of the rat type II glucocorticoid receptor cDNA, in the reverse orientation, downstream from the human neurofilament-L gene promoter. An antisense construct directed against the 3' non-coding region of the type II glucocorticoid receptor mRNA was used as this sequence contains less homology with other members of the steroid hormone superfamily of genes than other regions do and avoids cross-inhibition of other receptors. The plasmid (pUC 19) containing the human neurofilament-L gene promoter (pNF-L; gift from J.-P. Julien) was linearized by digestion with *Hind*III, which cleaves at a sequence at the 3' end of the promoter. Extruding sequences were filled in with the Klenow fragment of DNA polymerase and a blunt *Xba*I-*Pst*I fragment of the pSG-1 plasmid (gift from R. Miesfeld; this cDNA clone is a derivative of pRM16 (ref. 2)) was inserted into pNF-L in a reverse orientation. Finally, we added a *Pst*I-Bam HI fragment of the VP1 gene of SV40 which contains a polyadenylation signal.

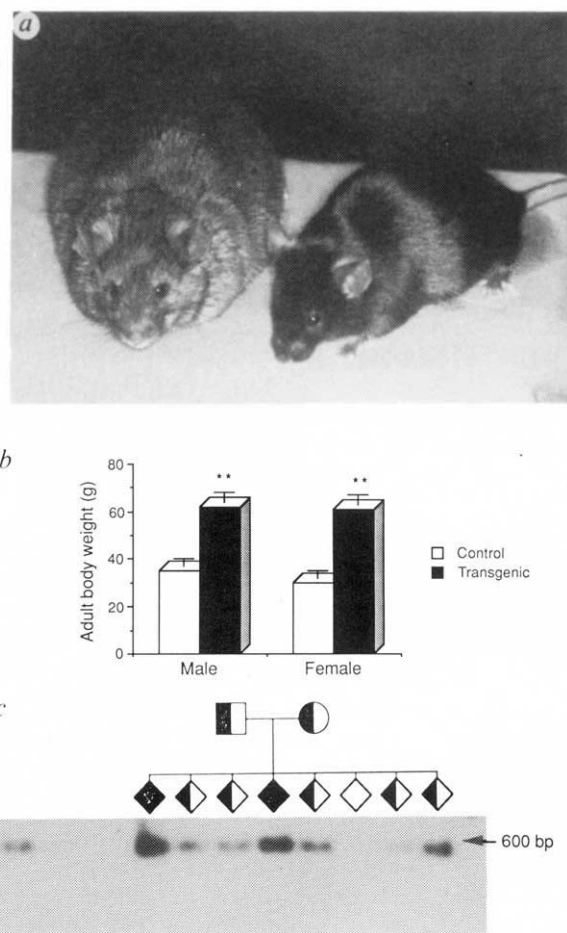


FIG. 2 Identification and appearance of transgenic mouse lines. **a**, Transgenic mice (transgenic mouse, line 1.3, left; normal mouse, right) were produced by microinjection of an ~4.9-kb *Nde*I-*Ase*I fragment excised from pNFL-AsGR plasmid (see Fig. 1) into the male pronucleus of fertilized mouse eggs (C3H×c57BL/6F2) obtained from superovulated prepubertal females 20 h after treatment with human chorionic gonadotropin²⁷. DNA from 13 liveborn animals was tested by Southern analysis for the presence of the insert and five animals were positive. **b**, The body weight of age-matched (6–7 months old) normal and transgenic animals ($n=8$) is shown. The significance of differences between means were tested by the Duncan-Kramer test²⁸ after analysis of variance. Two asterisks indicate $P < 0.01$. **c**, Southern blot analysis of the type II glucocorticoid receptor antisense transgene in littermates of F_1 generation heterozygotic parents derived from founder 1.3. Half-filled symbols represent heterozygotes, filled symbols are homozygotes and open symbols are nontransgenic offspring. A standard (extreme left lane), equivalent to 50 copies of the transgene, was also electrophoresed and hybridized with the probe. DNA (15 μ g per mouse) was digested with *Bam*HI, electrophoresed on 1.0% agarose gels, transferred to nylon filters (Hybond N⁺, Amersham), and hybridized in formamide buffer to a ³²P-labelled *Bam*HI-BamHI fragment containing a part of both the NF-L promoter and the antisense type II glucocorticoid receptor cDNA fragment. Homozygotic strains derived from 3 of the founder transgenic mice were used here.

the complete neurofilament-L gene, other than those of the 2.4-kb promoter region, may be necessary for absolute tissue-specific expression, and findings similar to ours, of less stringent neuronal expression directed by the 2.4-kb neurofilament promoter, have been described¹¹.

The transgene we incorporated into the mouse genome was designed to disrupt normal hypothalamus-pituitary-adrenal (HPA) axis function. The greatly increased plasma concentration of adrenocorticotrophic hormone (ACTH) (Table 1) of transgenic animals is suggestive of a failure of glucocorticoids to inhibit HPA axis activity at either, or both, central and pituitary sites. Plasma corticosterone concentration is also raised (Table 1), but not to the same extent as is ACTH, suggesting that the adrenal gland, which appears slightly hyperplastic, may become refractory to continuous stimulation. Glucocorticoids can decrease the number of glucocorticoid receptors in both hippocampus and pituitary gland^{12,13} and the high corticosterone levels in transgenic mice, secondary to decreased type II glucocorticoid receptor levels, could be an additional factor in downregulating the type II glucocorticoid receptor even further compared with antisense RNA.

Although we have induced HPA axis dysregulation in these animals, their most remarkable feature is greatly increased weight and fat deposition. Steroid hormones govern a large range of processes affecting development and physiological homeostasis and may influence feeding and behavioural patterns¹⁴. Glucocorticoids have a role in the development of obesity, caused by different aetiology^{15,16}. Our transgenic animals have a lowered maximum number of binding sites ($B_{\max}$) for glucocorticoids with raised serum corticosterone and ACTH concentrations, features also reported in the obese *fa/fa* Zucker rat¹⁷, although this has been questioned¹⁸.

Patients suffering from severe depression often show an increased activity of the HPA axis, as indicated by hypersecretion of corticotropin-releasing factor and cortisol, and a premature escape from the cortisol-suppressant action of

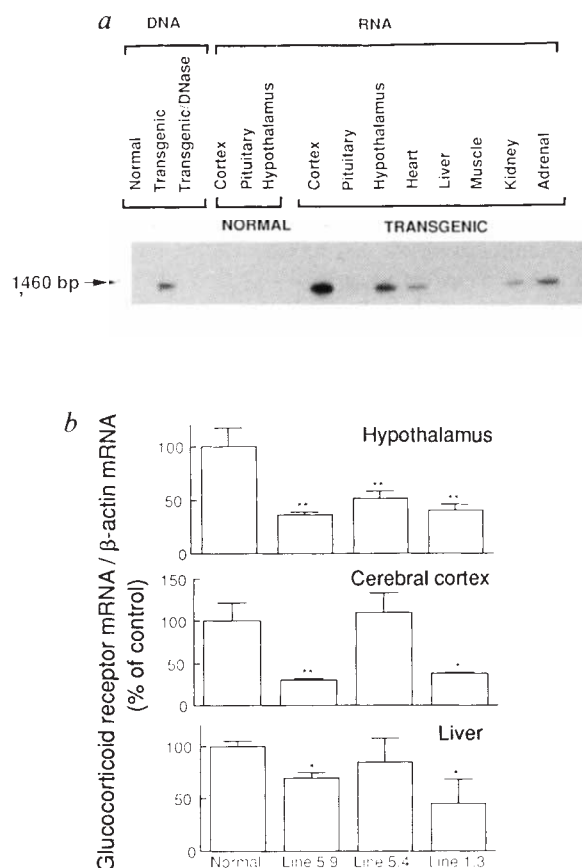
TABLE 1 HPA axis parameters in normal and transgenic mice

Glucocorticoid-receptor binding activity		
³ [H]dexamethasone bound (fmol per mg protein)		
Tissue	Normal	Transgenic line 1.3
Frontal cortex	18.50 ± 2.00 (n = 28)	11.90 ± 0.80 (n = 32)
Hypothalamus	5.70 ± 0.70 (n = 28)	4.00 ± 0.40 (n = 32)†
Kidney	3.30 ± 0.50 (n = 28)	2.30 ± 0.30 (n = 32)
Liver	0.65 ± 0.09 (n = 28)	0.45 ± 0.04 (n = 32)†
Pituitary gland	27.90 ± 3.00 (n = 12)	12.80 ± 0.40 (n = 16)*
Plasma hormone concentrations		
Animal	ACTH (pg ml ⁻¹ plasma)	Corticosterone (nM)
Normal	141 ± 35 (n = 6)	104 ± 19 (n = 19)
Transgenic line 1.3	716 ± 143 (n = 4)*	198 ± 43 (n = 14)†

Hormone content of blood plasma was measured by specific radioimmunoassay^{24,25}. Glucocorticoid binding activity in brain and other organs was measured with [³H]dexamethasone. Tissue was homogenized in 30 mM Tris, 1 mM EDTA, 10 mM molybdate, 10% (v/v) glycerol and 1 mM dithiothreitol (TEDGM, pH 7.4). After centrifugation at 105,000g for 45 min at 0 °C, an aliquot of the cytosol was incubated with 25 nM [³H]dexamethasone (specific activity 44.7 Ci per mmol; Amersham) for 20–24 h at 4 °C. The amount of nonspecific binding was determined in parallel incubations of the labelled steroid in the presence of a 500-fold excess of unlabelled RU 28362 (a type II glucocorticoid receptor-specific agonist; NEN-DUPONT, Boston). Sephadex-LH20 (Pharmacia) columns equilibrated with TEDGM buffer were used to separate bound from unbound steroid. After incubation, 100-μl aliquots were loaded onto the columns, washed with 100 μl TEDGM and eluted with 400 μl TEDGM into vials. The vials were then filled with 5 ml aqueous counting fluid (formula A-963; NEN) and counted in a LKB scintillation counter at 40% efficiency. Protein content was determined according to ref. 26. Differences between means was tested by the Duncan-Kramer test²⁸ after analysis of variance.

* $P < 0.01$; † $P < 0.05$.

FIG. 3 Transgene expression and effects on type II glucocorticoid receptor mRNA. *a*, Reverse transcriptase/PCR amplification assay of transgene expression. RNA, prepared by the guanidium/isothiocyanate method²⁹, was treated with enough DNase (10 U DNase/2.5 μg RNA for 10 min at 37 °C) to prevent amplification from DNA (third lane). A 15-base poly(dT) oligomer was used as primer for reverse transcriptase. A 21-base oligomer, corresponding to the sequence (1,824–1,844) immediately upstream from the polyadenylation signal in the VP1 gene of bacteriophage SV40, used in conjunction with a 20-base oligomer primer corresponding to positions 381–400 of the type II glucocorticoid receptor cDNA antisense fragment, served for PCR amplification of reverse transcriptase reaction products using a Perkin-Elmer Cetus PCR kit (Gene Amp). Conditions were as recommended by the manufacturer. After electrophoresis and transfer to Hybond N⁺, amplified products were identified by hybridization to a ³²P-labelled 20-base oligomer corresponding to positions 1,733–1,752 of the type II glucocorticoid receptor antisense cDNA. *b*, The type II glucocorticoid receptor mRNA/ β -actin mRNA ratio in extracts from hypothalamus (upper panel) cerebral cortex (middle panel) and liver (lower panel) was determined in normal mice and in 3 different lines of transgenic animals. RNA was separated on 0.8% agarose formaldehyde denaturing gels and blotted onto nylon filters (Hybond N, Amersham) before hybridization with type II glucocorticoid receptor cRNA antisense, type II glucocorticoid receptor cRNA sense and β -actin cRNA probes. Type II glucocorticoid receptor cRNA antisense probe was produced by T7 polymerase run-off transcription with [³²P]UTP of a 1.8-kb type II glucocorticoid receptor cDNA fragment² subcloned into the plasmid pGEM-1. Type II glucocorticoid receptor cRNA sense probe was generated from the same plasmid using SP6 polymerase. A β -actin cRNA probe was generated from a 1.5-kb β -actin cDNA *Pst*I fragment³⁰ inserted into pGEM-1. Hybridization conditions have been described⁵. Results are the mean ± s.e.m. of northern blot analyses from individual mice (normal, $n = 15$; transgenic lines 5.9, $n = 30$, 5.4 $n = 14$ and 1.3 $n = 10$). The statistical difference between means was calculated by the Duncan-Kramer test²⁸ after analysis of variance. * $P < 0.05$; ** $P < 0.01$.



dexamethasone^{19,20}. It is possible that this lack of sensitivity to glucocorticoids could be due to an abnormality of type II glucocorticoid receptor regulation at the level of the limbic-hypothalamic system, a hypothesis supported by the stimulatory effects of antidepressants on type II glucocorticoid receptor mRNA concentrations in brain^{21,22} and by their normalization of the HPA axis²³. Our transgenic animals with impaired type II glucocorticoid receptor function have increased HPA activity reminiscent of that in severe depression. These animals could thus be useful for studying obesity and the neuroendocrinological changes in depression, and for evaluating new antidepressant therapies. □

Received 31 July; accepted 18 November 1991.

1. Arriza, J. L. *et al.* *Science* **237**, 268–275 (1987).
2. Miesfeld, R. *et al.* *Nature* **312**, 779–781 (1984).
3. Ratka, A., Sutar, W., Bloemers, M. & de Kloet, E. R. *Neuroendocrinology* **50**, 117–123 (1989).
4. Meaney, M. J., Aitken, D. H., Viau, V., Sharma, S. & Sarrieu, A. *Neuroendocrinology* **50**, 597–604 (1989).
5. Pepin, M.-C. & Barden, N. *Molec. cell. Biol.* **11**, 1647–1653 (1991).
6. Katsuki, M. *et al.* *Science* **241**, 593–595 (1988).

7. Melton, D. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 144–148 (1985).
8. Kim, S. K. & Wold, B. J. *Cell* **42**, 129–138 (1985).
9. McGarry, T. J. & Lindquist, S. *Proc. natn. Acad. Sci. U.S.A.* **83**, 399–403 (1986).
10. Julien, J. P., Tretjakoff, I., Beaudet, L. & Peterson, A. *Genes Dev.* **1**, 1085–1095 (1987).
11. Julien, J.-P., Beaudet, L., Tretjakoff, I. & Peterson, A. *J. Physiol., Paris* **84**, 50–52 (1990).
12. Tornello, S., Orti, E., DeNicola, A. F., Rainbow, T. C. & McEwen, B. S. *Neuroendocrinology* **35**, 411–417 (1982).
13. Svec, F. & Rudis, M. *J. Biol. Chem.* **256**, 5984–5987 (1981).
14. Funder, J. W. & Sheppard, K. A. *Rev. Physiol.* **49**, 397–411 (1987).
15. Fletcher, J. M. *Br. J. Nutr.* **56**, 141–151 (1986).
16. Freedman, M. R., Horwitz, B. A. & Stern, J. S. *Am. J. Physiol.* **250**, R595–R607 (1986).
17. Douglas White, B. & Martin, R. J. *J. Steroid Biochem.* **36**, 681–686 (1990).
18. Langley, S. C. & York, D. A. *Brain Res.* **533**, 268–274 (1990).
19. Carroll, B. J., Curtis, G. C. & Mendels, J. *Arch. gen. Psychiat.* **33**, 1039–1050 (1976).
20. Linkowski, P. *et al.* *J. clin. Endocr. Metab.* **61**, 429–438 (1985).
21. Pepin, M.-C., Beaulieu, S. & Barden, N. *Molec. Brain Res.* **6**, 77–83 (1989).
22. Peiffer, A., Veilleux, S. & Barden, N. *Psychoneuroendocrinology* **16**, 505–515 (1991).
23. Greden, J. F. *et al.* *Arch. Gen. Psychiat.* **40**, 493–500 (1983).
24. Giguère, V., Côté, J. & Labrie, F. *Endocrinology* **110**, 1225–1230 (1982).
25. Provencher, P., Lorrain, A., Bélanger, A. & Fiet, F. *J. Steroid Biochem.* **36**, 589–596 (1990).
26. Bradford, M. M. *Analyt. Biochem.* **72**, 248–254 (1976).
27. Hogan, B., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo. A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1986).
28. Kramer, C. Y. *Biometrics* **12**, 307–310 (1956).
29. Chirgwin, J., Pryzbala, A., MacDonald, R. & Rutter, W. *Biochemistry* **8**, 5294–5299 (1979).
30. Farmer, S. R., Wan, K. M., Ben-Ze'ev, A. & Penman, S. *Molec. cell. Biol.* **3**, 182–189 (1983).

ACKNOWLEDGEMENTS. This work was supported by a Medical Research Council of Canada grant (N.B.) and FRSQ and MRC studentships (M.-C.P.).

Prevention of HIV-1 infection in chimpanzees by gp120 V3 domain-specific monoclonal antibody

E. A. Emini*, W. A. Schleif*, J. H. Nunberg*†, A. J. Conley*, Y. Eda‡, S. Tokiyoshi‡, S. D. Putney†§, S. Matsushita¶, K. E. Cobb||, C. M. Jett||, J. W. Eichberg†|| & K. K. Murthy||

* Merck, Sharp and Dohme Research Laboratories, West Point, Pennsylvania 19486, USA

‡ The Chemo-Sero-Therapeutic Research Institute, Kumamoto 861-15, Japan

§ Repligen Corporation, Cambridge, Massachusetts 02139, USA

¶ Kumamoto University, Kumamoto 860, Japan

|| Southwest Foundation for Biomedical Research, San Antonio, Texas 78228, USA

THE acquired immunodeficiency syndrome (AIDS) is the late-stage clinical manifestation of long-term persistent infection with the human immunodeficiency virus type 1 (HIV-1). Immune responses directed against the virus and against virus-infected cells during the persistent infection fail to mediate resolution of the infection. As a result, a successful AIDS vaccine must elicit an immune state that will prevent the establishment of the persistent infection following introduction of the virus into the host. The third hyper-variable (V3) domain of the HIV-1 gp120 envelope glycoprotein is a disulphide-linked closed loop of about 30 amino acids which binds and elicits anti-HIV-1 type-specific virus-neutralizing antibodies^{1–7}. The *in vitro* characteristics of anti-V3 domain antibody suggest that this antibody could by itself prevent HIV-1 infection *in vivo*^{8,9}, an idea supported by chimpanzee challenge studies in which protection against the HIV-1 persistent infection seemed to correlate with the presence of anti-V3 domain antibody^{10–12}. Here we directly demonstrate the protective efficacy of anti-V3 domain antibody *in vivo* and propose that this antibody is potentially useful as both a pre- and post-exposure prophylactic agent.

The antibody used for this purpose was a mouse-human IgG1 chimaeric monoclonal antibody, termed Cβ1, which contains

the intact variable region of the previously described murine 0.5β monoclonal antibody. The 0.5β antibody is directed to the V3 loop domain of the HIV-1 IIIb variant gp120 and has potent *in vitro* IIIb-specific virus-neutralizing activity¹³. The chimaeric Cβ1 antibody had identical *in vitro* specificity and potency (data not shown).

First, a purified preparation of Cβ1 was administered to a chimpanzee intravenously (i.v.) at a dose of 36 mg per kg body weight. The animal was challenged 24 h later with an i.v. administration of 75 chimpanzee-infectious doses of the HIV-1 IIIb variant. The animal had a circulating virus-neutralizing antibody titre of 1:320 at the time of challenge. An untreated control chimpanzee was simultaneously inoculated with virus.

Table 1 presents the results of the study. The control animal (x39) first presented signs of virus infection at three weeks post-challenge, based on virus isolation from its peripheral blood mononuclear cells (PBMCs). This animal specifically seroconverted by six weeks post-challenge and has remained antibody- and virus-isolation-positive during the 52-week observation period. In contrast, the animal treated with Cβ1 antibody (x289) has remained free of signs of virus infection. Attempts to detect virus-specific nucleic acid in the animal's PBMCs by polymerase chain reaction (PCR) have also been negative to date.

Following this initial observation of protection, a second study was started to assess the protective effect of Cβ1 when administered after virus challenge. Two chimpanzees (x299 and x196) were each inoculated i.v. with 75 chimpanzee-infectious doses of HIV-1 IIIb. Ten minutes post-inoculation, animal x299 was given Cβ1 antibody i.v. at a dose of 36 mg kg⁻¹. The total volume (~200 ml) was delivered in 30 min. The circulating neutralizing antibody titre at the end of Cβ1 administration was 1:640. Chimpanzee x196 was not treated.

The results are presented in Table 2. The control chimpanzee (x196) first yielded a positive virus isolation from PBMCs at six weeks post-challenge. The animal seroconverted at eight weeks and has remained virus-isolation- and antibody-positive during the 32-week observation period. The chimpanzee (x299) treated with Cβ1 antibody has not presented any signs of persistent virus infection to date. PCR analyses for virus-specific nucleic acid in the animal's PBMCs have also been negative.

These observations are direct evidence that anti-V3 loop, HIV-1-neutralizing antibody can, in the absence of other virus-specific immune responses, prevent an HIV-1 persistent infection. This establishes HIV-1 neutralizing activity mediated by the anti-V3 domain antibody as an *in vitro* correlate of an *in vivo* protective immune response. Although the Cβ1 antibody is specific for the IIIb variant of the virus, anti-V3 loop antibodies

† Present addresses: Genentech Inc., South San Francisco, California 94080, USA (J.H.N.); Alkermes Inc., Cambridge, Massachusetts 02139, USA (S.D.P.); Wyeth-Ayerst Research, Philadelphia, Pennsylvania 19087, USA (J.W.E.).

TABLE 1 Protective effect of pre-exposure administration of V3 loop-directed monoclonal antibody

Control chimpanzee (x39)					Pre-exposure C β 1-treated chimpanzee (x289)				
Weeks post-challenge	Virus-neutralizing antibody titre	Anti-HIV-1 ELISA and western blot	Virus isolation	Virus-specific PCR	Weeks post-challenge	Virus-neutralizing antibody titre	Anti-HIV-1 ELISA and western blot	Virus isolation	Virus-specific PCR
0	<10	—	—	—	0	<10	—	—	—
(1 day)	<10	—	ND	ND	(1 day)	320	+/-	ND	ND
1	<10	—	—	—	1	320	+/-	—	—
2	<10	—	—	—	2	80	+/-	—	—
3	<10	—	+	+	3	20	+/-	—	—
4	<10	—	+	+	4	10	—	—	—
6	<10	+	+	+	6	<10	—	—	—
8	20	+	+	+	8	<10	—	—	—
10	160	+	+	+	10	<10	—	—	—
12	320	+	+	+	12	<10	—	—	—
14	160	+	+	+	14	<10	—	—	—
16	≥ 640	+	+	+	16	<10	—	—	—
20	≥ 640	+	+	+	20	<10	—	—	—
24	≥ 640	+	+	+	24	<10	—	—	—
28	≥ 640	+	+	+	28	<10	—	—	—
32	≥ 640	+	+	+	32	<10	—	—	—
36	≥ 640	+	+	+	36	<10	—	—	—
40	≥ 640	+	+	+	40	<10	—	—	—
44	≥ 640	+	+	+	44	<10	—	—	—
48	≥ 640	+	+	+	48	<10	—	—	—
52	≥ 640	+	+	+	52	<10	—	—	—

C β 1 antibody was purified by protein A-affinity chromatography. Animals were challenged 1 day after administration of the antibody to chimpanzee x289. The challenge virus was kindly provided by Larry Arthur and Peter Fischinger (US National Cancer Institute) and was quantitated at the time of challenge as described by Emini *et al.*¹². The chimpanzees were maintained under conditions that met or exceeded the NIH guidelines for the care of laboratory animals. Virus-neutralizing antibody titre was measured in cell culture using the IIIb variant of HIV-1 as described by Robertson *et al.*¹⁵. The quantity of virus used in the assay was identical to the quantity used for chimpanzee challenge. The values represent the inverse of the highest serum dilution which yielded an effective virus neutralization as described. Anti-HIV-1 ELISA and western blot reactive antibodies were measured as described by Emini *et al.*¹². +, Positive enzyme-linked immunosorbent assay (ELISA) reaction and multiple reactive western blot bands; +/-, only gp120-reactive western blot band, negative ELISA; —, negative ELISA and no western blot reactive bands. Virus was isolated by *in vitro* culture of chimpanzee PBMCs using four methods. (1) Chimpanzee PBMCs were cocultured with an equal number of mitogen-activated human PBMCs in medium containing interleukin-2 (IL-2) and DEAE-dextran. (2) Chimpanzee PBMCs were activated by phytohaemagglutinin (PHA) and then cocultured as for method 1. (3) Chimpanzee PBMCs were stimulated with PHA and were then cultured alone in IL-2-containing medium. (4) Chimpanzee PBMCs were cocultured with an equal number of activated human PBMCs in medium supplemented with PHA and IL-2. In each case, cultures contained 1×10^7 chimpanzee cells in a volume of 10 ml. The cultures were incubated at 37 °C in a 5% CO₂ atmosphere for 4 weeks. Medium supernatants were assayed for HIV-1 p24 antigen production at weekly intervals. Virus growth by any of the methods was considered to be a positive virus isolation. The presence of HIV-1 specific nucleic acid in chimpanzee PBMCs was assayed by PCR. DNA was obtained from PBMCs by incubation of cell samples in 10 mM Tris HCl, pH 8.3, 1.0 mM EDTA, pH 8.0, 0.5% SDS, 150 μ g ml⁻¹ proteinase K at 60 °C for 60 min. After phenol/chloroform extraction, the DNA was precipitated by ethanol, and 1–3 μ g was used for the PCR. Each reaction mixture contained, in addition to the DNA, 1.5 mM MgCl₂, 10 mM Tris HCl, pH 8.3, 50 mM KCl, 0.032 mM of each dNTP, 50 ng of each primer and 1.0 unit *Taq* polymerase. The HIV-1 *gag*-specific primer pair SK38/39 was used. We used 35 reaction cycles: each cycle consisted of a 94 °C incubation for 1.5 min and a 56 °C incubation for 3.0 min. Amplified product was detected by hybridization with the SK19 probe. Sensitivity of the PCR assay was determined as 10 HIV-1-specific copies per DNA sample. ND, not done.

TABLE 2 Protective effect of post-exposure administration of V3 loop-directed monoclonal antibody

Control chimpanzee (x196)					Post-exposure C β 1-treated chimpanzee (x299)				
Weeks post-challenge	Virus-neutralizing antibody titre	Anti-HIV-1 ELISA and western blot	Virus isolation	Virus-specific PCR	Weeks post-challenge	Virus-neutralizing antibody titre	Anti-HIV-1 ELISA and western blot	Virus isolation	Virus-specific PCR
0	<10	—	—	—	0	640	+/-	—	—
(1 day)	<10	—	ND	ND	(1 day)	320	+/-	ND	ND
1	<10	—	—	—	1	320	+/-	—	—
2	<10	—	—	—	2	160	+/-	—	—
3	<10	—	—	—	3	80	+/-	—	—
4	<10	—	—	—	4	20	—	—	—
6	<10	—	+	+	6	10	—	—	—
8	<10	+	+	+	8	<10	—	—	—
10	<10	+	+	+	10	10	—	—	—
12	10	+	+	+	12	<10	—	—	—
14	20	+	+	+	14	<10	—	—	—
16	80	+	+	+	16	<10	—	—	—
18	80	+	+	+	18	<10	—	—	—
20	160	+	+	+	20	<10	—	—	—
24	320	+	+	+	24	<10	—	—	—
28	320	+	+	+	28	<10	—	—	—
32	320	+	+	+	32	<10	—	—	—

Techniques are described in the Table 1 legend. The day-0 antibody titre for chimpanzee x299 was determined using serum obtained immediately after C β 1 administration (see text). The virus-neutralizing antibody titre before C β 1 administration was <10. ND, not done.

can be elicited with broader neutralizing activity¹⁴. Our results support efforts to develop vaccine immunogens designed to induce such broadly reactive anti-V3 loop antibodies. They also indicate that an appropriate broadly reactive anti-V3 loop monoclonal antibody may serve as an effective pre-exposure or post-exposure immunoprophylactic agent. □

Received 8 October; accepted 21 November 1991.

- Leonard, C. K. *et al. J. biol. Chem.* **265**, 10373–10382 (1990).
- Goudsmit, J. *et al. AIDS* **2**, 157–164 (1988).
- Goudsmit, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 4478–4482 (1988).
- Ho, D. D. *et al. J. Virol.* **61**, 2024–2028 (1987).
- Javaherian, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 6768–6772 (1989).
- Kenealy, W. R. *et al. AIDS Res. Human Retroviruses* **5**, 173–182 (1989).
- Rusche, J. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 3198–3202 (1988).
- Linsley, P. S., Ledbetter, J. A., Kinney-Thomas, E. & Hu, S.-L. *J. Virol.* **62**, 3695–3702 (1988).
- Skinner, M. A. *et al. J. Virol.* **62**, 4195–4200 (1988).
- Berman, P. W. *et al. Nature* **345**, 622–625 (1990).
- Girard, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 542–546 (1991).
- Emini, E. A. *et al. J. Virol.* **64**, 3674–3678 (1990).
- Matsushita, S. *et al. J. Virol.* **62**, 2107–2114 (1988).
- Robertson, G. A., Kostek, B. M., Schleif, W. A., Lewis, J. A. & Emini, E. A. *J. virol. Meth.* **20**, 195–202 (1988).
- Javaherian, K. *et al. Science* **250**, 1590–1593 (1990).

ACKNOWLEDGEMENTS. We thank D. Titus, A. Rhodes, S. Rouse and G. I. Melendez for technical assistance, J. Lewis for discussions, and D. Wilson for preparation of the manuscript.

Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm

Serafín Piñol-Roma & Gideon Dreyfuss*

Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6148, USA

RNA polymerase II transcripts, heterogeneous nuclear RNAs (hnRNAs), associate in the nucleus with specific proteins that bind premessenger RNA (hnRNP proteins)^{1,2} and with small nuclear ribonucleoprotein particles (snRNPs)^{3–5}. These hnRNA–hnRNP–snRNP complexes assemble on nascent transcripts and hnRNA is processed to mRNA in them^{6–8}. hnRNP proteins have been localized to the nucleoplasm^{9–13} and their functions were presumed to be limited to nuclear events in mRNA biogenesis. It was proposed that an exchange of hnRNP for mRNA-binding proteins accompanies transport of mRNA from the nucleus to the cytoplasm^{1,14}. We show here that several of the abundant hnRNP proteins, including A1, shuttle between the nucleus and the cytoplasm. hnRNP proteins may thus also have cytoplasmic functions. Furthermore, when in the cytoplasm, A1 is bound to mRNA and RNA polymerase II transcription is necessary before it can return to the nucleus. We propose that the cytoplasmic ribonucleoprotein complex of mRNA with hnRNP proteins is the substrate of nuclear–cytoplasmic transport of mRNA.

hnRNP complexes in vertebrate cells are composed of at least 20 main proteins^{15,16}. Previous studies have shown that hnRNP proteins that disperse during mitosis reaccumulate in the newly formed daughter cell nuclei by two modes: transcription-independent (for example, C proteins) or transcription-dependent (for example, A1 protein)¹⁷. To test whether the nuclear localization of A1 during interphase is also dependent on RNA polymerase II (pol II) transcription, we exposed HeLa cells to a pol II inhibitor, actinomycin D^{18,19}, for three hours. The intracellular location of the hnRNP C proteins (relative molecular mass (M_r) 41,000/43,000 by SDS-PAGE) and A1 protein (M_r 36,000) was then determined by immunofluorescence microscopy using monoclonal antibodies^{12,16}. As expected,

if pol II transcription is necessary for the accumulation of A1 in the nucleus, A1 accumulated in the cytoplasm of cells treated with actinomycin D ($5 \mu\text{g ml}^{-1}$) (Fig. 1Ad). This cytoplasmic staining was not apparent before¹⁷ owing to the shorter incubation time with the inhibitor (45 min). In contrast, C proteins remain restricted to the nucleus in actinomycin D-treated cells (Fig. 1Ac), arguing that our treatment does not cause the nuclei to become generally leaky. Surprisingly, protein synthesis inhibitors do not detectably diminish the cytoplasmic signal of

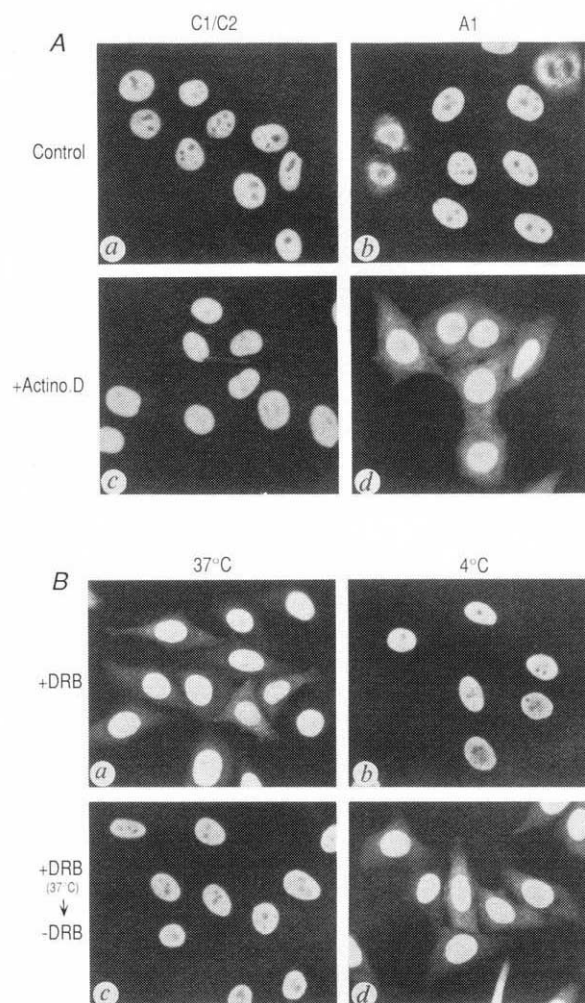


FIG. 1 A, hnRNP A1 protein accumulates in the cytoplasm of cells treated with $5 \mu\text{g ml}^{-1}$ actinomycin D. HeLa cells grown on glass coverslips were incubated for 3 h in the absence (Control; a and b) or presence (+actino.D; c and d) of actinomycin D at $5 \mu\text{g ml}^{-1}$. Cells were then fixed and immunostained with the anti-C-protein monoclonal antibody 4F4 (ref. 12) (a, c) or the anti-A1 monoclonal antibody 4B10 (ref. 16) (b and d). Treatment of cells with actinomycin D results in cytoplasmic accumulation of A1, whereas the C proteins remain sequestered in the nucleus. The pair of cells with some cytoplasmic staining (b) are recently divided daughter cells in which A1 has still not fully reaccumulated in the nucleus (see ref. 17). B, hnRNP A1 accumulated in the cytoplasm in the presence of transcriptional inhibitors is pre-existing A1 exported from the nucleus. a, b, Immunofluorescence staining with 4B10 (a) and phase microscopy view of the same field (b) of HeLa cells incubated for 3 h with $5 \mu\text{g ml}^{-1}$ actinomycin D in the presence of $20 \mu\text{g ml}^{-1}$ cycloheximide. Inhibition of new A1 protein synthesis does not diminish the cytoplasmic signal observed with 4B10 in actinomycin D-treated cells (compare 1Ad). c, d, Immunofluorescence staining with 4B10 (c) and phase microscopy view (d) of the same field of cells incubated with $20 \mu\text{g ml}^{-1}$ cycloheximide in the absence of actinomycin D. The distribution of A1 in these cells is indistinguishable from that observed without inhibitors (compare 1Ab). Cycloheximide alone has no effect also on the nuclear localization of the C proteins (data not shown). Indirect immunofluorescence staining has been described^{11,12}.

* To whom correspondence should be addressed.

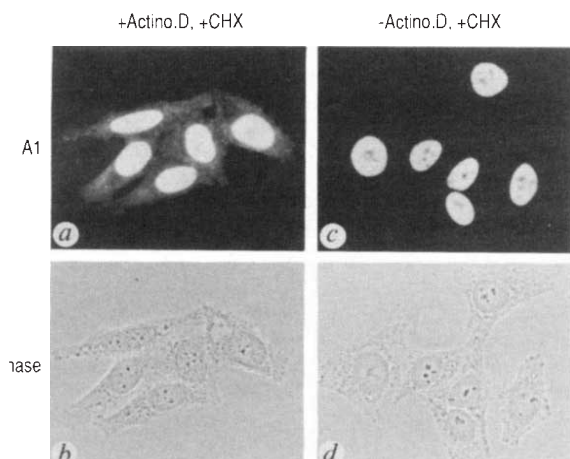


FIG. 2 Reversibility and temperature-dependence of the cytoplasmic accumulation of A1 in cells treated with DRB. *a*, HeLa cells were incubated in the presence of 100 μ M 5,6-dichlororibofuranosyl benzimidazole (DRB) for 3 h at 37 °C and then fixed and immunostained with 4B10. *b*, HeLa cells treated like those in *a*, except that incubation with DRB was at 4 °C. *c*, Cells were incubated with DRB at 37 °C for 3 h. Culture medium was then replaced with fresh medium without the inhibitor and the cells incubated at 37 °C for another hour before fixation and staining with 4B10. *d*, Cells treated as *c*, except that they were incubated in DRB-free medium at 4 °C.

A1 observed in the presence of actinomycin (Fig. 1*B*). Therefore, A1 that accumulates in the cytoplasm after treatment with pol II inhibitors is pre-existing A1 that was exported from the nucleus. In addition to A1, other hnRNP proteins such as A2 and E also accumulate in the cytoplasm in the presence of pol II inhibitors, whereas several others, including protein U, appear to be restricted to the nucleus (S.P.R., H. Siomi and G.D., unpublished results). No effect on the intracellular distribution

of either A1 (Fig. 1*Bc*) or C proteins was found with translational inhibitors alone, or with actinomycin D at 0.04 μ g ml⁻¹ (data not shown), a concentration that inhibits pol I but not pol II transcription¹⁹.

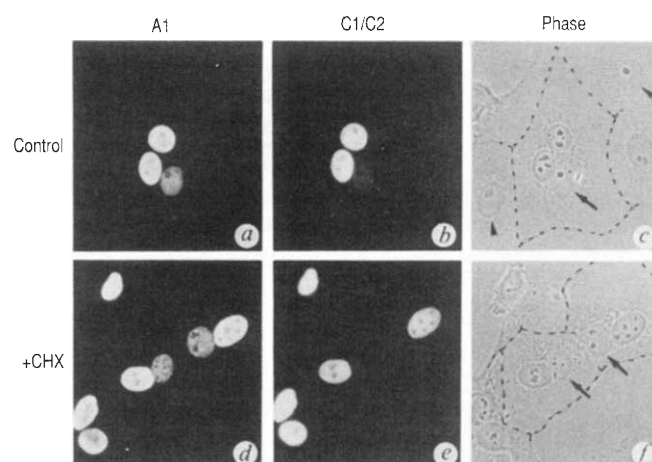
Similar results were observed with 5,6-dichlororibofuranosyl benzimidazole (DRB), a pol II-transcription inhibitor that is structurally and mechanistically unrelated to actinomycin D^{20,21} (Fig. 2*a*). A1 failed to accumulate in the cytoplasm however, when DRB treatment was carried out at 4 °C (Fig. 2*b*), suggesting that export of A1 from the nucleus is an active transport. Furthermore, the cytoplasmic accumulation of A1 is readily reversed if transcription is restored by placing DRB-treated cells into fresh DRB-free medium at 37 °C (Fig. 2*c*) but not at 4 °C (Fig. 2*d*). In all of these experiments the C proteins are detected exclusively in the nucleus (data not shown).

These observations indicate that A1 shuttles between the nucleus and the cytoplasm. To establish whether this occurs in the absence of transcriptional inhibitors, we tested the ability of HeLa A1 and C proteins to migrate between nuclei in human/*Xenopus laevis* interspecies heterokaryons by direct, double-label immunofluorescence staining. This approach has been used to study the shuttling of nucleolar proteins between the nucleolus and the cytoplasm²². Figure 3 shows that, 4 hours after fusion, HeLa A1 is present in both HeLa and *X. laevis* nuclei. This is also observed in the presence of protein-synthesis inhibitors (Fig. 3*d*), showing that the human A1 protein found in the *X. laevis* nuclei was mostly pre-existing A1 which originated from the human nucleus in the heterokaryons. But human C proteins appear to be restricted to the nucleus as they are detected in low amounts in the *X. laevis* nuclei of heterokaryons only when new protein synthesis is allowed (Fig. 3*b,e*). Therefore, hnRNP A1 (but not C proteins) also shuttles between the nucleus and the cytoplasm in the absence of inhibitors of transcription.

To determine whether A1 is bound to mRNA when it is in the cytoplasm, we exposed untreated cells and cells treated with

FIG. 3 hnRNP A1 protein migrates between nuclei in interspecies heterokaryons but C proteins do not. *a–c*, HeLa and *X. laevis* culture cells grown on glass coverslips were fused with polyethylene glycol. Four hours after fusion, cells were fixed and immunostained simultaneously with TRITC (tetramethylrhodamine isothiocyanate)-labelled 4B10 and fluorescein isothiocyanate-labelled 4F4. The distribution of A1 (*a*) and C proteins (*b*) was examined in a Zeiss Axiophot microscope using the rhodamine or fluorescein channel, respectively. *d–f*, HeLa-*X. laevis* heterokaryons were generated as in *a–c*, except that cycloheximide was added to the culture medium to a concentration of 20 μ g ml⁻¹ for 30 min before fusion, and this concentration was maintained throughout culture of the heterokaryons. *d*, Localization of A1 and *e*, localization of C proteins. *c* and *f*, Phase contrast microscopy view of the corresponding fields. *X. laevis* nuclei in heterokaryons are indicated by arrows, and in unfused cells by arrowheads. For clarity, the outline of the heterokaryons is indicated with a dashed line. Neither 4F4 nor 4B10 stain *X. laevis* nuclei that are not in heterokaryons, whereas all HeLa nuclei are intensely stained with these antibodies. All heterokaryons show similar patterns of staining. Human A1 can be detected in *X. laevis* nuclei of heterokaryons within 2 h of fusion (not shown).

METHODS. HeLa-*X. laevis* heterokaryons were produced by a modification of the procedure described in ref. 22. Briefly, *X. laevis* cells were grown on glass coverslips to subconfluent density. Trypsinized HeLa cells were plated onto the *X. laevis*-bearing coverslips, and allowed to attach for 2–4 h in heterokaryon growth medium (DME, 10% fetal calf serum, 20% water, 1% penicillin and streptomycin). Each coverslip was rinsed in PBS and inverted onto a drop of 50% polyethylene glycol (PEG; Gibco BRL) pre-warmed to 29 °C for 2 min. The coverslip was then rinsed twice in PBS to remove PEG and placed into a 35 mm culture plate in heterokaryon growth medium. Heterokaryons were incubated at 29 °C under 5% CO₂ for 4 h before fixation. Where indicated, cycloheximide (20 μ g ml⁻¹) was added to the HeLa-*X. laevis* cell co-culture at 30 min before fusion, and incubation continued in the presence of cycloheximide. Cells were fixed and immunostained with FITC-labelled 4F4 and TRITC-labelled 4B10. The monoclonal antibodies 4F4 (ref. 12) and 4B10 (ref. 16) against hnRNP C and A1 proteins, respectively, were purified from ascites fluid and labelled with the fluorophores as described¹⁷.



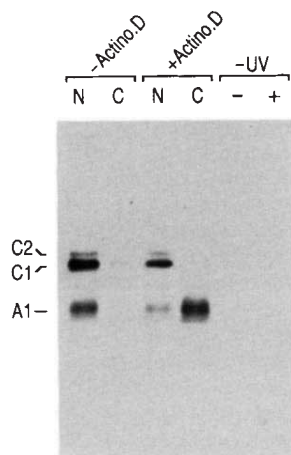


FIG. 4 HnRNP A1 accumulated in the cytoplasm of cells treated with pol II inhibitors is bound to poly(A)⁺ RNA (mRNA). HeLa cells incubated in the absence (lanes -Actino.D) or presence (lanes +Actino.D) of 5 $\mu\text{g ml}^{-1}$ actinomycin D for 3 h were exposed to ultraviolet light to induce covalent RNA-protein crosslinks *in vivo*^{14,23}. Poly(A)⁺ RNA-protein crosslinked complexes were isolated from the cytoplasmic and nuclear fractions (lanes 'C' and 'N' respectively) by chromatography on oligo(dT), in the presence of SDS and β -mercaptoethanol. Under these protein-denaturing conditions, only proteins that were bound to RNA during irradiation with ultraviolet light co-purify with the RNA. Proteins crosslinked to RNA were released from the complexes by digestion with RNase, resolved by SDS-PAGE, and the C and A1 proteins detected by immunoblotting with antibodies 4F4 and 4B10, respectively. Lanes labelled '-UV': cells incubated in the absence (-) or presence (+) of actinomycin D, but the ultraviolet irradiation step was omitted. Extraction of RNA, affinity chromatography on oligo(dT), and immunoblotting were carried out as described for the samples irradiated with ultraviolet light, except that the nucleocytoplasmic fractionation steps were omitted.

METHODS. Photochemical RNA-protein crosslinking by irradiation with ultraviolet light of cells on culture dishes, and isolation and analysis of RNPs was carried out as before^{14,23}. For the nuclear-cytoplasmic fractionation step, cells were lysed in RSB-100 (containing 0.5% Triton X-100, 0.5% aprotinin, 1 $\mu\text{g ml}^{-1}$ each of leupeptin and pepstatin A, and 10 mM vanadyladenosine) by four passages through a 25-gauge syringe needle. Fractions were separated by brief centrifugation at 3,000g. Material from equivalent cell numbers was loaded in each lane.

actinomycin D (5 $\mu\text{g ml}^{-1}$) to ultraviolet light to induce covalent protein-RNA crosslinks *in vivo*^{14,23}. Proteins crosslinked to poly(A)⁺ RNA were then isolated from the nuclear and cytoplasmic fractions and analysed by SDS-PAGE and immunoblotting. As expected, both the A1 and C proteins are crosslinked almost exclusively to nuclear poly(A)⁺ RNA in untreated cells (Fig. 4). In cells treated with actinomycin D, however, both A1 and C proteins are detected in the nuclear poly(A)⁺ RNA-containing fraction and a large amount of A1 (but not C proteins) is

crosslinked to cytoplasmic poly(A)⁺ RNA. We found a reproducible and considerable decrease in the amount of A1 and C proteins crosslinked to nuclear RNA in actinomycin D-treated cells, which is consistent with a reduction in the amount of poly(A)⁺ RNA in the nuclei of these cells, presumably as mRNA is transported to the cytoplasm. Neither A1 nor C proteins co-purify with the RNA if the ultraviolet irradiation is omitted (lanes labelled minus UV in Fig. 4). Thus some, if not all, of the A1 that accumulates in the cytoplasm in cells treated with pol II inhibitors is bound to poly(A)⁺ RNA.

The shuttling of hnRNP proteins and their presence in the cytoplasm indicate that these proteins may have functions outside the nucleus in cytoplasmic mRNA metabolism. Several nucleolar proteins are shuttled between the nucleus and the cytoplasm (see ref. 22), and this process could have a role in nucleocytoplasmic transport of ribosomal components and in the cytoplasmic regulation of nuclear activity. The cytoplasmic hnRNP-mRNA complex is a novel RNP and it may be a cytoplasmic precursor to the mRNP-mRNA complex, the predominant cytoplasmic form of mRNA. If so, an exchange of hnRNP for mRNP proteins in the cytoplasm would be another step for post-transcriptional regulation. The cytoplasmic accumulation of A1 bound to poly(A)⁺ RNA in the absence of pol II transcription is probably a result of the trapping of this RNP in the cytoplasm, because pol II transcription is required for the nuclear import of A1, its accumulation in the nucleus or its removal from the mRNA in the cytoplasm.

Intron-containing pre-mRNAs seem to be spliced as nascent transcripts^{6,24} and any remaining introns that can form spliceosomes retain unspliced pre-mRNA in the nucleus^{25,26}. Presumably proper 3'-end processing is also required before mRNA can be transported to the cytoplasm. hnRNP proteins restricted the nucleus (such as the C proteins) may serve to retain the incompletely processed pre-mRNA in the nucleus, and need to be selectively removed from the mRNA before or during mRNA translocation to the cytoplasm. The 5'-cap structure facilitates, but is not sufficient for, the export of pol II transcripts²⁷. The finding of A1 bound to poly(A)⁺ RNA both in the nucleus and in the cytoplasm raises the possibility that mRNA may be exported from the nucleus bound to the shuttling hnRNP proteins. Nuclear export of 5S RNA is possibly mediated by the proteins L5 and/or TFIIB²⁸, and the Balbiani ring mRNAs in *Chironomus tentans* salivary glands travel through nuclear pores as ribonucleoprotein particles²⁹⁻³¹. Given our findings, we propose that the hnRNP-mRNA complex is the export substrate. If hnRNP proteins are bound to mRNA during its transit through the nuclear pores, the signal for the export may reside on the mRNA, in which case hnRNP proteins may simply be exported 'piggy-backed' on the mRNA. Alternatively, and we consider it likely, hnRNP proteins may play a more active part in mRNA export, operating as a specialized hnRNP protein translocation process in which mRNA is exported as a result of its association with the shuttling hnRNP proteins. □

Received 9 October; accepted 25 November 1991.

- Dreyfuss, G. A. *Rev. Cell Biol.* **2**, 459-498 (1986).
- Dreyfuss, G., Swanson, M. S. & Piñol-Roma, S. *Trends Biochem. Sci.* **13**, 86-91 (1988).
- Steitz, J. A. *et al.* in *Functions of the Abundant snRNPs: Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles* (ed. Birnstiel, M. L.) (Springer, Berlin, 1988).
- Lührmann, R., Kastner, B. & Bach, M. *Biochim. biophys. Acta* **1087**, 265-292 (1990).
- Gall, J. G. *Science* **252**, 1499-1500 (1991).
- Osheim, Y. N., Miller, O. L. & Beyer, A. L. *Cell* **43**, 143-151 (1985).
- Fakan, S., Leser, G. & Martin, T. E. *J. Cell Biol.* **103**, 1153-1157 (1986).
- Wu, Z., Murphy, C., Callan, H. G. & Gall, J. G. *J. Cell Biol.* **113**, 465-483 (1991).
- Martin, T. E. & Okamura, C. S. in *The Cell Nucleus* (ed. Busch, H.) Vol. 9 119-144 (Academic, New York, 1981).
- Leser, G. P., Escara-Wilke, J. & Martin, T. E. *J. Biol. Chem.* **259**, 1827-1833 (1984).
- Dreyfuss, G., Choi, Y. D. & Adam, S. A. *Molec. cell. Biol.* **4**, 1104-1114 (1984).
- Choi, Y. D. & Dreyfuss, G. *J. Cell Biol.* **99**, 1997-2004 (1984).
- Piñol-Roma, S., Swanson, M. S., Gall, J. G. & Dreyfuss, G. *J. Cell Biol.* **109**, 2575-2587 (1989).
- Dreyfuss, G., Adam, S. A. & Choi, Y. D. *Molec. cell. Biol.* **4**, 415-423 (1984).
- Choi, Y. D. & Dreyfuss, G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7471-7475 (1984).
- Piñol-Roma, S., Choi, Y. D., Matunis, M. J. & Dreyfuss, G. *Genes Dev.* **2**, 215-227 (1988).

- Piñol-Roma, S. & Dreyfuss, G. *Science* **253**, 312-314 (1991).
- Hurwitz, J., Furth, J. J., Malamy, M. & Alexander, M. *Proc. natn. Acad. Sci. U.S.A.* **48**, 1222-1230 (1962).
- Perry, R. P. & Kelley, D. E. *J. Cell Physiol.* **76**, 127-140 (1970).
- Tamm, I. & Seghal, P. B. *Adv. Virus Res.* **22**, 187-258 (1978).
- Egyhazi, E., Pigon, A. & Rydlander, L. *Eur. J. Biochem.* **122**, 445-451 (1982).
- Borer, R. A., Lehner, C. F., Eppenberger, H. M. & Nigg, E. A. *Cell* **56**, 379-390 (1989).
- Adam, S. A., Choi, Y. D. & Dreyfuss, G. *J. Virol.* **57**, 614-622 (1986).
- Beyer, A. L. & Osheim, Y. N. *Genes Dev.* **2**, 754-765 (1988).
- Chang, D. D. & Sharp, P. A. *Cell* **59**, 789-795 (1989).
- Legrain, P. & Rosbash, M. *Cell* **57**, 573-583 (1989).
- Hamm, J. & Mattaj, J. W. *Cell* **63**, 109-118 (1990).
- Guddat, U., Bakken, A. H. & Pieler, T. *Cell* **60**, 619-628 (1990).
- Stevens, B. J. & Swift, H. *J. Cell Biol.* **31**, 55-78 (1966).
- Skoglund, U., Anderson, K., Björkroth, B., Lamb, M. M. & Daneholt, B. *Cell* **34**, 847-855 (1983).
- Mehlin, H., Skoglund, U. & Daneholt, B. *Expt. Cell Res.* **133**, 72-77 (1991).

ACKNOWLEDGEMENTS. We thank the members of our laboratory for discussion and for comments on the manuscript. This work was supported by the NIH and the Howard Hughes Medical Institute.

Dynamin is a GTPase stimulated to high levels of activity by microtubules

Howard S. Shpetner & Richard B. Vallee

Cell Biology Group, Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts 01545, USA

DYNAMIN was initially identified in calf brain tissue as a protein of relative molecular mass 100,000 which induced nucleotide-sensitive bundling of microtubules¹. Purified dynamin showed only trace ATPase activity. But in combination with an activating factor removed during the purification, it exhibited microtubule-activated ATPase activity and dynamin-induced bundles showed evidence of ATP-dependent force production¹. Dynamin is the product of the *Drosophila* gene *shibire*^{2,3}, which has been implicated in synaptic vesicle recycling^{4,5} and, more generally, in the budding of endocytic vesicles from the plasma membrane^{6,7}. Dynamin also shows⁸ extensive homology with proteins that participate in vacuolar protein sorting⁹ and spindle pole-body separation¹⁰ in yeast, and in interferon-induced viral resistance^{11,12} in mammals. All members of this family contain consensus sequence elements consistent with GTP binding near their amino termini, although none has been shown to have GTPase activity. We report here that dynamin is a specific GTPase which can be stimulated to very high levels of activity by microtubules.

Earlier, we observed microtubule-stimulated dynamin ATPase activity upon reconstitution of the purified protein with an activating fraction (denoted S1')¹ removed during purification in a microtubule-cosedimentation step. As the dynamin polypeptide (M_r 100K) contains primary sequence elements consistent with GTP binding⁸, we evaluated the substrate specificity of the purified protein. Dynamin hydrolysed GTP at a rate of $126 \text{ nmol min}^{-1} \text{ mg}^{-1}$ (0.20 s^{-1}) (Table 1), whereas ATP, CTP, UTP and GDP were all hydrolysed at low or undetectable rates.

TABLE 1 Substrate specificity of the dynamin nucleotidase

Nucleotide	Hydrolysis rate ($\text{nmol min}^{-1} \text{ mg}^{-1}$)	
	(-) Tubulin	(+) Tubulin
Mg^{2+} -GTP	126 ± 4	$1,979 \pm 187$
Mg^{2+} -ATP	≤ 2	4 ± 2
Mg^{2+} -CTP	5 ± 2	≤ 2
Mg^{2+} -UTP	13 ± 2	13 ± 4
Mg^{2+} -GDP	≤ 2	≤ 2

Dynamin nucleotide-hydrolysis rates were assayed at 37°C in PEM/G buffer (5 mM PIPES, pH 7.0, 50 mM sodium glutamate, 1 mM EGTA, 1 mM MgSO_4 , 1 mM dithiothreitol) by colorimetric measurement of phosphate release³³. Dynamin was purified as described²⁴, except that PEM/G buffer replaced the phosphate-containing buffer (P/G buffer) used after the initial microtubule extraction step. Microtubules were assembled from DEAE-purified tubulin²⁵ using taxol, washed²⁴, and resuspended in PEM/G buffer. Assay mixtures were prepared as described²⁴. Errors indicate range of duplicate data points. Nucleotide concentration, 0.5 mM. Microtubule concentration, 0.15 mg ml^{-1} .

Microtubules activated the rate of GTP hydrolysis 16-fold, up to $1,979 \text{ nmol min}^{-1} \text{ ml}^{-1}$ (3.13 s^{-1}). In other experiments (Figs 1 and 2), maximal rates of microtubule-stimulated hydrolysis ranged from 925 to $3,837 \text{ nmol min}^{-1} \text{ mg}^{-1}$ (1.46 – 6.06 s^{-1} ; Figs 1 and 2). Both basal and microtubule-activated GTPase activities copurified with dynamin during DEAE-Sepharose chromatography (Fig. 1a).

GTP hydrolysis exhibited a simple Michaelis-Menten dependence on GTP, with a K_m of $12 \pm 2 \mu\text{M}$ ($r^2 = 0.997$), which is below reported physiological levels (50 – $150 \mu\text{M}$)¹³. At $100 \mu\text{M}$ GTP, microtubule-activated hydrolysis was not inhibited by physiological levels of ATP (0.5 – 2.5 mM)¹⁴; Table 2). Microtubule-activated hydrolysis was inhibited by AlF_3 -S, GTP- γ -S, GMP-PCP, GMP-PNP and *N*-ethylmaleimide (NEM), but not by GDP, NaVO_3 , NaN_3 , or AMP-PNP at the concentrations assayed (Table 2).

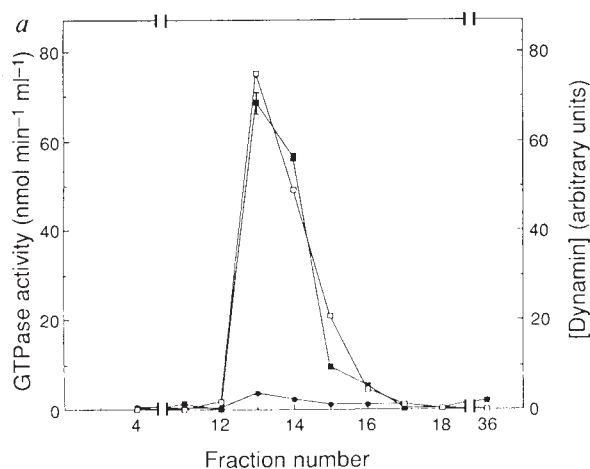
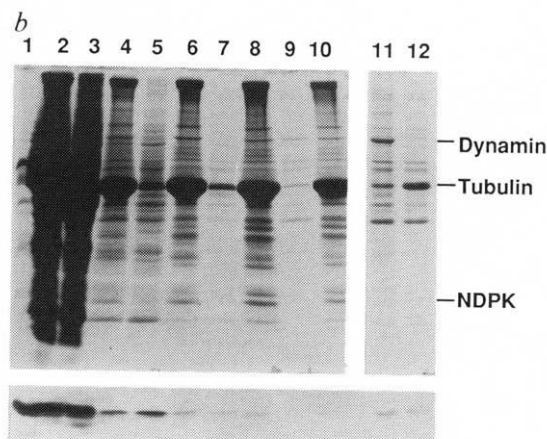


FIG. 1 a, Copurification of dynamin with basal and microtubule-activated GTPase activities. Dynamin was extracted from bovine brain microtubules using GTP and AMP-PNP, rebound to microtubules, and re-extracted with GTP and salt. The GTP-salt extract (S2') was desalted and fractionated on a DEAE-Sepharose column²⁴. Column fractions were desalted and assayed for GTPase activity in the presence (■) or absence (●) of microtubules (0.11 mg ml^{-1}). 48% of the total microtubule-stimulated activity applied to the column was recovered in the column fractions. The peak microtubule-stimulated activity was $3,837 \text{ nmol min}^{-1} \text{ mg}^{-1}$. The relative concentration of dynamin (□) was determined by densitometry of an 8% polyacrylamide SDS gel of the column fractions (not shown; compare ref. 1, Fig. 3a and ref. 24, Fig. 4). For details of GTPase assays, see legend to Table 1. b, Fractionation of NDPK immunoreactivity during dynamin purification. An 8–18%



gradient polyacrylamide-SDS gel³⁵ (top) and immunoblot³⁶ (bottom) was probed with an antibody against a peptide from the NDPK encoded by the murine nm23 gene (Ab. 11 ref. 37; gift from L. Liotta). Lane 1, beef liver NDPK (Boehringer Mannheim); lane 2, bovine brain white matter cytosol; lanes 3, 4, supernatant and pellet after taxol-assisted microtubule assembly; lanes 5, 6, supernatant and pellet after first microtubule buffer wash; lanes 7, 8, supernatant and pellet after second microtubule buffer wash; lanes 9, 10: supernatant (E') and pellet after extraction of microtubules with GTP and AMP-PNP, lane 11, E' extract at fivefold higher concentration than in lane 9, lane 12, S1' supernatant (dynamin activating fraction) after centrifugation of E' with microtubules. 30 μl of all samples were loaded on the gel. Samples in lanes 2–10 were resuspended to the starting volume of cytosol.

As the dynamin nucleotidase was highly specific for GTP, we reasoned that the apparent microtubule-activated ATPase activity observed with the activating fraction (S1')¹ could have been due to hydrolysis of GTP, generated from ATP and trace levels of GDP (expected in S1') by nucleoside diphosphokinase (NDPK) activity. We did find that the dynamin activating fraction (S1') exhibited an NDPK activity¹⁵ of 57 ± 24 nmol min⁻¹ ml⁻¹ ($n = 3$), sufficiently high to account for the observed dynamin ATPase activity (up to 12 nmol min⁻¹ ml⁻¹)¹. In addition, low levels of NDPK were detected in this fraction by immunoblotting (Fig. 1b, lane 12).

Recent evidence has indicated a role for NDPK in mitosis and in determining the metastatic potential of tumour cells^{16,17}. Microtubules have been suggested as a possible target of NDPK, in view of biochemical¹⁷⁻¹⁹ and immunocytochemical¹⁷ evidence for an association with the enzyme. But NDPK clearly did not copurify quantitatively with either microtubules or dynamin in our preparations (Fig. 1b), and a role for NDPK in dynamin function is uncertain.

We performed several experiments to evaluate the specificity with which microtubules stimulated dynamin GTPase activity. Activity increased with microtubule concentration up to a maximum level at $50\text{--}80$ $\mu\text{g ml}^{-1}$ (Fig. 2a), somewhat lower than that required (1 mg ml⁻¹) for maximal activation of either kinesin^{20,21} or cytoplasmic dynein²² ATPase activity, and within the physiological concentration range of microtubules. GTPase activity did not exhibit a simple Michaelis-Menten dependence on microtubule concentration, and in some experiments a decrease in activity was noted at higher microtubule concentrations. These properties could reflect the complex and apparently cooperative interactions of dynamin with microtubules¹. Tubulin subunits (0.13 mg ml⁻¹) stimulated the dynamin GTPase to only 9% of the extent seen with the same concentration of microtubules. Purified MAP 2 strongly inhibited microtubule activation of GTPase activity (Fig. 2b), suggesting that dynamin might interact with the acidic C termini of the tubulin subunits, as

TABLE 2 Pharmacological characterization of dynamin GTPase activity

Inhibitor	GTPase activity (% control)
0.5 mM GTP- γ -S	10 ± 1
0.5 mM GMP-PNP	37 ± 11
0.5 mM GMP-PCP	9 ± 1
0.5 mM GDP	137 ± 16
2.0 mM ATP*	94 ± 2
0.5 mM AMP-PNP	117 ± 16
2.4 mM NaF†	101 ± 5
2.4 mM NaF, 0.1 mM AlCl ₃ †	23 ± 4
0.2 mM NEM†,‡	44 ± 3
1 mM NaN ₃ †	92 ± 5
1 mM NaVO ₃ †	108 ± 5

GTPase activities were measured by either colorimetric³³ or isotopic³⁴ measurement of phosphate release. For details of protein preparation and assay conditions, see legend to Table 1. GTP concentration was 0.25 mM unless otherwise indicated. Microtubule concentration, $0.10\text{--}0.14$ mg ml⁻¹. Errors indicate range or standard deviations of 2-4 experiments.

* [GTP] = 0.10 mM.

† [GTP] = 1.0 mM.

‡ Samples of dynamin were incubated with NEM for 10 min at room temperature. Dithiothreitol was then added to 20 mM, and the samples kept on ice until assay. Control samples contained 20 mM DTT without NEM.

previously proposed for cytoplasmic dynein²³. Other anionic biopolymers—actin and vimentin filaments, and polyglutamic acid (PGA)—did not appreciably alter dynamin GTPase activity at the concentrations assayed (Fig. 2b). In view of recent genetic evidence implicating dynamin in endocytosis^{2,3}, we also evaluated the effect of reconstituted clathrin coats on GTPase activity. No significant effect was observed (Fig. 2b). Thus, activation of the dynamin GTPase appeared to be specific for microtubules.

Tubulin is, itself, a GTPase. Thus, it seemed plausible that stimulation of GTPase activity could represent dynamin activa-

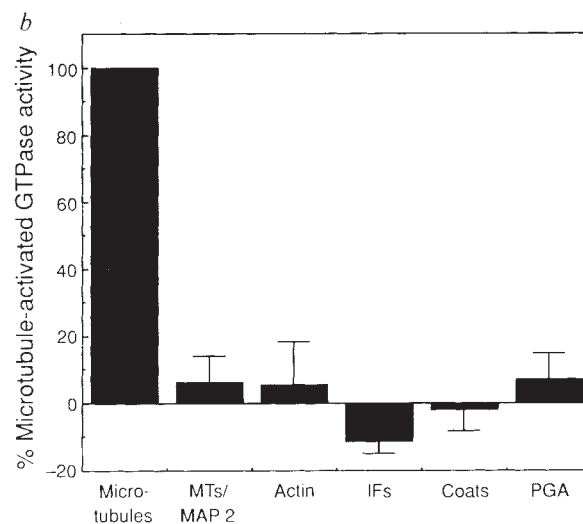
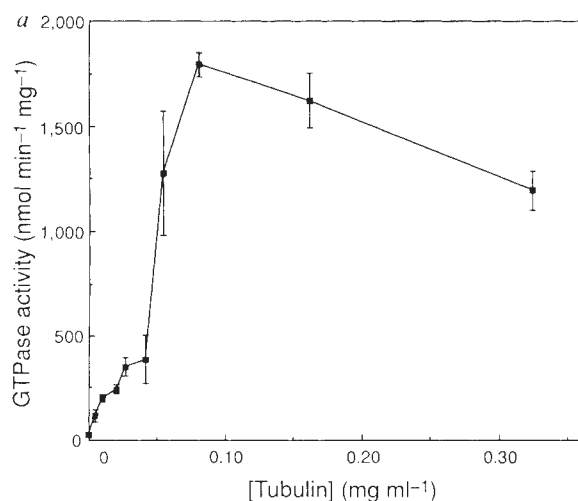


FIG. 2 Effect of microtubules and other polymers on dynamin GTPase activity. *a*, Dependence of GTPase activity on microtubule concentration. *b*, Relative effects of microtubules and other anionic polymers on dynamin GTPase activity.

METHODS. For details of protein preparation and assay conditions, see legend to Table 1. GTP concentration was 1 mM. For *a*, GTPase activities were determined by isotopic³⁴ measurement of phosphate release; error reported as in Table 1. In *b*, GTPase activities were determined by colorimetric³³ measurement of phosphate release. Purified F-actin³⁸, vimentin intermediate filaments (IFs)³⁹, reconstituted clathrin coats⁴⁰, polyglutamic acid (PGA; Sigma), gel filtration-purified MAP 2 (ref. 25) or microtubules^{24,25} were either resuspended or dialysed into PEM/G buffer and assayed for

their effect on dynamin GTPase activity. The effect of MAP2 on GTPase activity was determined in the presence of microtubules. The activities of dynamin and the polymers alone were subtracted from the total activities, and the data expressed as a percentage of the microtubule-activated GTPase activity. Results shown are means $\pm$ s.d. of either four (actin, PGA) or two (coats, IFs, MAP 2) experiments. Activities of dynamin alone ranged from 6.9 to 19.1% of microtubule-stimulated activities, which ranged from 925 to $3,546$ nmol min⁻¹ mg⁻¹. Actin = $0.12\text{--}0.18$ mg ml⁻¹; IFs = 0.05 mg ml⁻¹; coats = $0.14\text{--}0.17$ mg ml⁻¹; PGA = $0.08\text{--}0.24$ mg ml⁻¹; MAP2 = 0.10 mg ml⁻¹, microtubules = $0.10\text{--}0.14$ mg ml⁻¹, except for comparison with IFs where microtubules = 0.05 mg ml⁻¹.

tion of the tubulin GTPase. To address this question, we prepared microtubules²⁴ from DEAE-purified tubulin²⁵ labelled by incorporation of [α -³²P]GTP (1.46 ± 0.23 mole GTP incorporated per mole tubulin) during polymerization with taxol. The microtubules (0.44 mg ml^{-1}) were then incubated in either the presence or absence of dynamin ($10 \mu\text{g ml}^{-1}$) in an excess of cold GTP (1 mM) at 37°C for 20 min and centrifuged. The fraction of bound radiolabel found in the supernatant was independent of the presence of dynamin ($23 \pm 3\%$ with dynamin, $20 \pm 2\%$ without). Thus, guanine nucleotide exchange by microtubules was unaffected by dynamin.

The effect of microtubules on the dynamin GTPase reported here is comparable with their effect on the kinesin^{20,21} and cytoplasmic dynein²² ATPases. Indeed, the turnover number exhibited by dynamin with microtubules is equal to or higher than the microtubule activated ATPase activities of kinesin^{20,21} and cytoplasmic dynein²² and close to the actin-activated ATPase activity of myosin^{26,27}. It is also far greater than the GTP hydrolysis rates (k_{cat}) of G proteins ($2\text{--}4 \text{ min}^{-1}$)²⁸ and of Ras proteins in the presence of GTPase activating protein (GAP) (0.4 min^{-1})²⁸. Thus, the characteristics of the dynamin GTPase seem to be more consistent with a mechanochemical than a regulatory role.

The *shibire*^{ts} phenotype has revealed a role for *Drosophila* dynamin in endocytic membrane budding, a process not known to involve microtubules directly. The *shibire*^{ts} phenotype could indirectly reflect the interaction of late endosomes with microtubules^{29,30}. But it is also possible that, in our *in vitro* assays, microtubules mimic an as-yet-unidentified physiological activator of dynamin, or even tubulin in some novel membrane-associated form. The precise role of dynamin in vesicle budding remains to be determined: as noted above, it could have a mechanochemical role, perhaps even in vesicle closure, or it could serve a regulatory role, as proposed for the small GTP-binding proteins^{31,32}. Further work will be needed to distinguish between these possibilities. □

Genetic linkage of Werner's syndrome to five markers on chromosome 8

Makoto Goto*, Mark Rubenstein†, James Weber‡, Kathryn Woods§ & Dennis Drayna†||

† Departments of Molecular Biology and § Scientific Computing, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

* Department of Rheumatology, Tokyo Metropolitan Otsuka Hospital, 2-8-1 Minami-Otsuka, Tokyo 170, Japan

‡ Marshfield Foundation for Medical Research, 1000 N. Oak Street, Marshfield, Wisconsin 54449, USA

|| To whom correspondence should be addressed

WERNER'S syndrome (WS) is a rare autosomal recessive disease in which the affected individuals display symptoms of premature ageing¹⁻³. The substantial phenotypic overlap between WS and normal ageing indicates that these two conditions may have pathogenetic mechanisms in common³⁻⁵. The WS mutation has pleiotropic effects, and patients and their cells show many differences compared with normals⁵. Despite extensive study of the clinical and biochemical features of this disorder, the primary genetic defect remains unknown. We have undertaken a genetic linkage study in an effort to identify the locus of the primary defect⁶. Here we report close genetic linkage of the WS mutation to a group of markers on chromosome 8.

We studied 21 Japanese families⁷⁻¹⁶. These families originated in 16 different prefectures, and apart from an increased prevalence in Ibaragi and its neighbouring prefectures, were widely distributed throughout Japan. Thirteen of these families contained consanguineous marriages (mostly first cousins), and two others were suspected to be consanguineous. Eleven of the

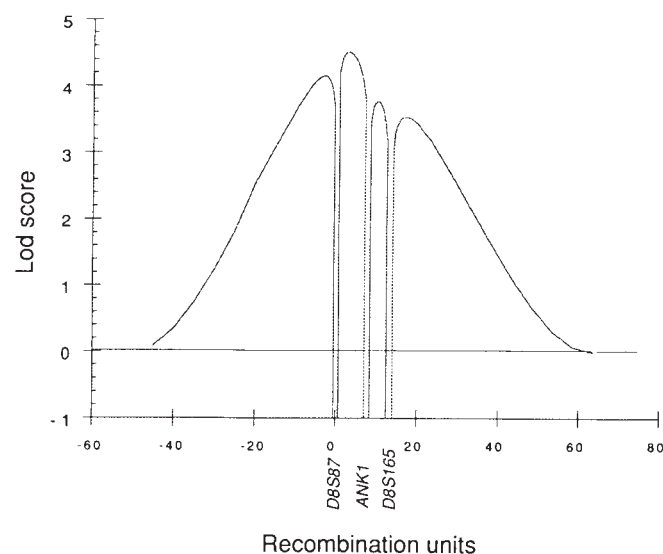


FIG. 1 Results of LINKMAP analysis of Werner's syndrome gene location. The locations of the marker loci were fixed as follows: $D8S87 \leftarrow 10 \text{ units} \rightarrow ANK1 \leftarrow 5 \text{ units} \rightarrow D8S165$, with the $D8S87$ locus set as zero. The distances between marker loci are those obtained from three-point ILINK analysis (analysing the marker loci only) in our Werner's syndrome two-generation families. Similarly, the marker order was the most likely given by three point ILINK. But the most likely order of $D8S87-ANK1-D8S165$ was only three times that of the second most likely order $D8S87-D8S165-ANK1$. The lod score was evaluated in 10 increments per interval. Analysis excluded families represented by only a single inbred affected individual.

Received 18 October; accepted 21 November 1991.

- Shpetner, H. S. & Vallee, R. B. *Cell* **59**, 421-432 (1989).
- Chen, M. S. *et al. Nature* **351**, 583-586 (1991).
- van der Bliek, A. M. & Meyerowitz, E. M. *Nature* **351**, 411-414 (1991).
- Poodry, C. A. & Edgar, L. J. *Cell Biol.* **81**, 520-527 (1979).
- Koenig, J. H., Kosaka, T. & Ikeda, K. *J. Neurosci.* **9**, 1937-1942 (1989).
- Kosada, T. & Ikeda, K. *J. Cell Biol.* **97**, 499-507 (1983).
- Kessel, I., Holst, B. D. & Roth, T. F. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4968-4972 (1989).
- Obar, R., Collins, C. A., Hammarback, J. A., Shpetner, H. S. & Vallee, R. B. *Nature* **347**, 256-261 (1990).
- Rothman, J. H., Raymond, C. K., Gilbert, T., O'Hara, P. J. & Stevens, T. H. *Cell* **61**, 1063-1074 (1990).
- Yeh, E., Driscoll, R., Coltrera, M., Olins, A. & Bloom, K. *Nature* **349**, 713-715 (1991).
- Staeheli, P. & Sutcliffe, J. G. *Molec. cell Biol.* **8**, 4524-4528 (1988).
- Staeheli, P., Haller, O., Boll, W., Lindermann, J. & Weissmann, C. *Cell* **44**, 147-158 (1986).
- de S. Otero, A. *Biochem. Pharmac.* **39**, 1399-1404 (1990).
- Henderson, J. F. & Paterson, A. R. P. *Nucleotide Metabolism* 18-22 (Academic, New York, 1973).
- Agarwal, R. P., Robison, B. & Parks, R. E. *Jr Meth. Enzym.* **51**, 376-386 (1978).
- Steeg, P. S. *et al. J. natn. Cancer Inst.* **80**, 200-204 (1988).
- Biggs, J., Hersperger, E., Steeg, P. S., Liotta, L. A. & Shearn, A. *Cell* **63**, 933-940 (1990).
- Penningroth, S. M. & Kirschner, M. W. *J. molec. Biol.* **115**, 643-673 (1977).
- Burns, R. G. & Islam, K. B. *Eur. J. Biochem.* **117**, 515-519 (1981).
- Wagner, M. C., Pfister, K. C., Bloom, G. S. & Brady, S. T. *Cell Motil. Cytoskel.* **12**, 195-215 (1989).
- Kuznetsov, S. A. & Gelfand, V. I. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8530-8534 (1986).
- Shpetner, H. S., Paschal, B. M. & Vallee, R. B. *J. Cell Biol.* **107**, 1001-1009 (1988).
- Paschal, B. M., Obar, R. A. & Vallee, R. B. *Nature* **342**, 569-572 (1989).
- Shpetner, H. S. & Vallee, R. B. *Meth. Enzym.* **196**, 181-192 (1991).
- Vallee, R. B. *Meth. Enzym.* **134**, 104-116 (1986).
- Pollard, T. D., Doberstein, S. K. & Zot, H. G. *A. Rev. Physiol.* **53**, 653-681 (1991).
- Harrington, W. R. & Rodgers, M. E. *A. Rev. Biochem.* **53**, 35-73 (1984).
- Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **348**, 125-132 (1990).
- Matteioni, R. & Kreis, T. E. *J. Cell Biol.* **105**, 1253-1265 (1987).
- Herman, B. & Albertini, D. F. *J. Cell Biol.* **98**, 565-576 (1984).
- Balch, W. E. *Trends Biochem. Sci.* **15**, 473-477 (1990).
- Hall, A. *Science* **249**, 635-640 (1990).
- Lanzetta, P. A., Alvarez, L. J., Reinach, P. S. & Candia, O. A. *Analyt. Biochem.* **100**, 95-97 (1979).
- Paschal, B. M., Shpetner, H. S. & Vallee, R. B. *Meth. Enzym.* **196**, 181-192 (1991).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Bloom, G. S. & Vallee, R. B. *J. Cell Biol.* **96**, 1523-1531 (1986).
- Rosengard, A. *et al. Nature* **342**, 177-180 (1989).
- Vidulich, J. A. & Watt, S. J. *biol. Chem.* **246**, 4866-4871 (1971).
- Spikstrom, K. L., Miller, R. K. & Goldman, R. D. *Meth. Enzym.* **196**, 506-525 (1991).
- Keen, J. H., Willingham, M. C. & Pastan, I. H. *Cell* **16**, 303-312 (1979).

ACKNOWLEDGEMENTS. We thank Yu-li Wang, R. Goldman, C. Burgess, and B. Paschal for their respective gifts of G-actin, bovine lens vimentin, coated vesicles and MAP2.

families consisted of a single generation, nine had two generations, and one had three generations. Seven families contained several affected siblings. The sample consisted of 63 individuals, 31 of which were affected (for details of the pedigrees, see Supplementary Information). Heparinized venous blood was obtained from all individuals with informed consent. Permanent cell lines were established by Epstein-Barr transformation of lymphocytes¹⁷, or by simian virus-40 transformation of fibroblasts¹⁸.

Werner's syndrome was diagnosed using standard clinical criteria^{3,7,19}, which include: (1) characteristic habitus and stature; (2) premature senescence; (3) scleroderma-like skin changes, and (4) endocrine abnormalities. At least three of these four major signs were required for the diagnosis. Diagnosis was confirmed by demonstration of hyaluronuria in eight individuals²⁰ and of diminished lifespan of cultured skin fibroblasts in 20 individuals^{18,21}. Individuals were assigned as unaffected only if they were more than 30 years old.

The marker loci used in our primary survey of the human genome consisted predominantly of cytosine-adenine-repeat microsatellite markers²², of which 128 were from the Marshfield collection²² and a further 23 were from the literature. In addition, the marker set also included a few trinucleotide (TTA-repeat) and tetranucleotide (TTTA- and TCTA-repeat) markers. Markers were typed by polymerase chain reaction (PCR) followed by polyacrylamide gel resolution of the PCR products²².

We tested a total of 156 polymorphic markers for linkage to Werner's syndrome. These markers were distributed on all of the autosomes, with a median heterozygosity of 65% in the Japanese population. These markers had a 90% probability of detecting linkage (at 10 centimorgans or less) to ~60% of the autosomal component of the human genome.

Our first indication of linkage appeared when we noted that two chromosome-8 markers, *D8S87* (Mfd39) and *ANK1* (ankyrin), showed increased homozygosity among affected individuals from inbred families. Such increases in homozygosity are predicted for markers that lie close to a rare autosomal locus²³. Traditional maximum likelihood calculations using data from all the families confirmed linkage for these two loci and also revealed linkage to several other nearby markers. The markers and their characteristics in the Japanese population are listed in Table 1; the genotypes in the WS families are available as Supplementary Information.

Significantly positive two-point lod scores were obtained for linkage between WS and five loci that lie on chromosome 8 (refs 24, 25). The two-point scores at several different values of θ are listed in Table 2. There was evidence for close linkage between WS and the two markers *D8S87* and *ANK1*. Evidence for linkage was also obtained for three apparently more distant markers, *D8S165*, *D8S166* and *D8S164*, although the pairwise lod scores for these linkages varied considerably. *LPL* (lipoprotein lipase),

Locus	Marker	Alleles*	Frequency†	Observed heterozygosity†
<i>ANK1</i>	ankyrin	107	0.63	0.45
		113	0.37	
<i>D8S87</i>	Mfd39	145	0.23	0.60
		149	0.01	
		151	0.57	
		153	0.06	
		155	0.13	
<i>D8S165</i>	Mfd117	142	0.28	0.57
		146	0.57	
		148	0.01	
		150	0.01	
		152	0.09	
<i>D8S166</i>	Mfd159	154	0.04	0.87
		112	0.02	
		114	0.02	
		116	0.32	
		118	0.03	
		120	0.01	
		122	0.06	
		124	0.17	
		126	0.16	
		128	0.07	
<i>D8S164</i>	Mfd104	130	0.04	0.84
		132	0.07	
		134	0.03	
		165	0.29	
		167	0.05	
		171	0.09	
		173	0.08	
		175	0.12	
		189	0.04	
		191	0.22	
		193	0.02	
		195	0.02	
		197	0.01	
		199	0.06	

* Alleles are the size in nucleotides of the predominant PCR product.

† Allele frequencies and the observed heterozygosity were determined in 50 unrelated Japanese individuals.

another marker on 8p, showed no linkage to WS in our data set. Iterative two-point calculations using the MLINK program²⁶ gave $\theta_{\max}$ values for these six markers as listed in Table 2.

Although none of the pairwise lod scores exceeded 5, multi-locus analysis gave significantly larger lod scores. Using the ILINK program²⁶, the cumulative lod score for linkage of *D8S87*-WS-*ANK1* was 7.69. Finally, using the data set that excluded families represented by just a single affected individual, we obtained a four-point lod score for the order *D8S87*-WS-

FIG. 2 Genotypes of recombinant inbred WS individuals. Individuals are offspring of first-cousin marriages. Numbers beneath each individual are the alleles of the markers listed; 'W' denotes the allele containing the mutation at the Werner's syndrome locus. For each individual, the arrows indicate the interval in which a recombination has occurred at a meiosis somewhere between the affected individual and his or her parents' common ancestor. Solid lines denote portions of the chromosomes that are homozygous; dashed lines denote portions which, as a result of recombination, are heterozygous. *LPL* lies distal to *ANK1* (ref. 33). These individuals support an order of: pter-*LPL*-(*D8S87*, WS)-*ANK1*-*D8S165*-*D8S166*-*D8S164*-cen.

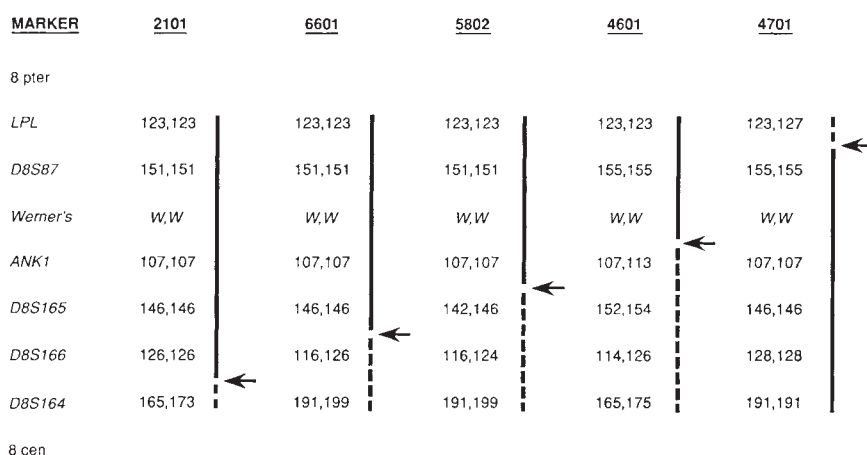


TABLE 2 Lod scores

Pairwise lod scores*										
Marker										
<i>D8S87</i>	−∞	5.1	4.8	4.1	3.3	2.5	1.8	1.2	0.63	0.23
<i>ANK1</i>	−∞	2.9	2.7	2.3	1.9	1.4	1.0	0.64	0.34	0.12
<i>D8S165</i>	−∞	1.5	1.9	1.8	1.5	1.2	0.81	0.46	0.18	0.02
<i>D8S166</i>	−∞	2.5	3.5	3.5	3.1	2.4	1.9	1.3	0.75	0.31
<i>D8S164</i>	−∞	−4.6	−1.6	−0.37	0.19	0.39	0.39	0.28	0.13	0.01
<i>LPL</i>	−∞	−2.9	−1.6	−0.92	−0.56	−0.34	−0.21	−0.12	−0.07	−0.03
Recombination fraction (θ)										
	0.0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
Maximum pairwise lod scores†										
Marker	θ	Lod								
<i>WS-D8S87</i>	0.050	5.14								
<i>WS-ANK1</i>	0.058	2.89								
<i>WS-D8S165</i>	0.11	1.91								
<i>WS-D8S166</i>	0.12	3.57								
<i>WS-D8S164</i>	0.28	0.45								
<i>WS-LPL</i>	0.50	0.00								

* Lod scores obtained in two-point analysis, using the MLINK program of the LINKAGE v5.1 program package, obtained from the Utah anonymous ftp site 'corona.med.utah.edu'²⁶ run on a VAX 9000 computer, rounded to two significant figures. Values are obtained after breaking loops in the consanguineous families. The values for marker loci *D8S164* and *D8S166* were obtained on an IBM-PC computer using LINKAGE v5.03 from J. Ott (Columbia University) run on individual families and then added. The data for these two markers were analysed using only the alleles observed in each family, plus an additional allele that represented the collective frequency of all alleles that did not occur in that family. The allele frequencies used were those in the normal Japanese population.

† Maximum pairwise lod scores obtained as already described, using recombination increments of 0.001 units.

ANK1-D8S165 of 9.92. We also examined the support for gene order in our data set. Using three-point analysis, the most likely order of the three closely linked loci was *D8S87-WS-ANK1*. But this order was only 2.19 times more likely than the second most likely order of *WS-D8S87-ANK1*, and 7.6 times more likely than the order *WS-ANK1-D8S87*.

We also used the program LINKMAP (ref. 26) to determine the most likely location of the gene for WS in this region. This analysis shows that the most likely location of the WS locus lies between *D8S87* and *ANK1*, roughly five recombination units from each. (Fig. 1). Several other points in this region give lod scores within 1 lod of the maximum, and thus the 90% confidence limits of this location are broadly based on this data.

More data on the order of these loci were obtained by examination of the genotypes of affected individuals from consanguineous marriages. In these individuals, homozygosity considerations allowed us to examine the genotypes of the occasional inbred individuals who are heterozygous for closely linked markers²³. As shown in Fig. 2, these genotypes suggest that the order of these loci is 8pter-*LPL*-(*D8S87, WS*)-*ANK1-D8S165-D8S166-D8S164*-cen. It is noteworthy that six inbred affected individuals (individuals 1701, 4601, 4801, 6001, 6701 and 6801) are heterozygous for either *ANK1* or *D8S87*. However, no inbred affecteds are heterozygous for both. This result supports the idea that these two marker loci flank the locus for WS.

This order is supported by results from others who have found the order 8pter-*D8S87-ANK1*-8cen (ref. 27) and 8pter-*D8S87-D8S165-D8S166-D8S164*-8cen (J.W., unpublished result). Although evidence for the order of the WS locus is not large, these results suggest the order of loci in this region is 8pter-*LPL-D8S87-WS-ANK1-D8S165-D8S166-D8S164*-8cen.

Werner's syndrome is a clinically complex disease, in that virtually all tissues of the body seem to be affected. The genetics of this disorder, however, argue for a relatively simple, single gene mutation as the cause of these diverse symptoms. For example, the disease is rare and displays a clear autosomal recessive mode of inheritance^{3,7}. In addition, Werner's syndrome is highly penetrant, and the WS phenotype is quite consistent among the affected individuals. Our linkages clustered on chromosome 8 are in agreement with the idea of a mutation at a single locus as the cause of this disease.

The closest marker, *D8S87*, is known to be tightly linked to the gene for tissue plasminogen activator, which lies on chromosome 8, band p12 (ref. 28). Another apparently close marker, *ANK1*, has been independently localized to the same chromosomal region, 8p11.1-21.1 (refs 24, 29). Although the other markers have been physically localized only to chromosome 8 (J.W., unpublished results), these data strongly suggest that the WS mutation lies close to or within the region 8p12.

Werner's syndrome has been reported worldwide^{1,3,30-32}, and although the clinical phenotype of Werner's syndrome is similar in all populations, it is conceivable that the WS phenotype could be conferred by different mutations in other, non-Japanese populations. This possibility can now be quickly evaluated using these closely linked markers. These markers can be used for presymptomatic, prenatal and carrier diagnosis in WS families of Japanese and perhaps other ancestries. They may also be useful to evaluate the hypothesis that other progeroid syndromes, such as Hutchinson-Guillford progeria and Rothmund-Thompson syndrome, may be due to different mutations at this same locus^{2,5}. The closest markers also provide a starting point for efforts to clone the gene that is defective in WS. It is possible that more detailed analysis of the Werner's syndrome mutation may lead to a greater understanding of the factor or factors that determine the characteristic lifespan in man. □

Received 5 August; accepted 6 December 1991.

- Werner, O. thesis, Univ. Kiel (1904).
- Thannhauser, S. J. *Ann. Int. Med.* **23**, 559 (1945).
- Epstein, C. J., Martin, G. M., Schultz, A. L. & Motulsky, A.G. *Medicine* **45**, 177-221 (1966).
- Salk, D. *Hum. Genet.* **62**, 1-15 (1982).
- Werner's Syndrome and Human Aging (eds Salk, D., Fujiwara, Y. & Martin, G. M.) (Plenum, New York, 1985).
- Botstein, D., White, R., Skolnick, M. & Davis, R. *Am. J. hum. Genet.* **32**, 314-331 (1989).
- Goto, M., Tanimoto, K., Horiuchi, Y. & Sasazuki, T. *Clin. Genet.* **19**, 8-15 (1981).
- Matsunaga, I., Iwata, S. & Yoshida, T. *Jpn. J. Ophthalmol.* **69**, 743 (1975).
- Chi, K. et al. *West. Jpn. J. Dermat.* **50**, 824-827 (1988).
- Aoki, R. & Gon, H. *Jpn. J. Med.* **77**, 125 (1988).
- Mori, S. et al. *Jpn. J. Geront.* **25**, 486-490 (1988).
- Fujisawa, A., Shoji, J., Terada, H. & Kitano, S. *Jpn. J. Ophthalmol.* **31**, 6785-689 (1989).
- Mashiyama, S., Mizoi, K., Takahashi, A., Suzuki, J. & Saito, T. *Jpn. J. Neurosurg.* **16**, 75-78 (1988).
- Hayashi, N., Kiura, Y., Aoki, M. & Tamaki, T. *Jpn. J. Orthop.* **31**, 1871-1873 (1985).
- Okuno, N. et al. *Diabetes* **33**, 179 (1990).
- Shindo, Y., Akiyama, J., Matsumoto, K., Takase, Y. & Hashimoto, T. *J. Dermat.* **13**, 396-398 (1986).
- Thorley-Lawson, D. A., Chess, L. & Strominger, J. L. *J. exp. Med.* **146**, 495-508 (1977).
- Matsumura, T., Nagata, M., Konishi, R. & Goto, M. in *Advances in Experimental Medicine and Biology: Werner's Syndrome and Human Aging* Vol. 190 (eds Salk, D., Fujiwara, Y. & Martin, G. M.) 313-330 (Plenum, New York, 1985).

19. Goto, M. in *Neurocutaneous Diseases* (ed. Gomez, M.R.) 242–246 (Butterworths, Boston, 1987).
20. Goto, M. & Murata, K. *Clin. chim. Acta* **85**, 101–106 (1978).
21. Hanaoka, F. *et al.* in *Advances in Experimental Medicine and Biology: Werner's Syndrome and Human Aging* Vol. 190 (eds Salk, D., Fujiwara, Y. & Martin, G.M.) 439–457 (Plenum, New York, 1985).
22. Weber, J. L. & May, P. E. *Am. J. hum. Genet.* **44**, 388–396 (1989).
23. Lander, E. S. & Botstein, D. *Science* **236**, 1567–1570 (1987).
24. Kitatani, M., Chiyo, H., Ozaki, M., Shike, S. & Miwa, S. *Hum. Genet.* **78**, 94–95 (1988).
25. Weber, J. L., Kwitek, A. E., May, P. E., Patterson, D. & Drabkin, H. *Nucleic Acids Res.* **18**, 4038 (1990).
26. Lathrop, G. M. & Lalouel, J.-M. *Am. J. hum. Genet.* **36**, 460–465 (1984).
27. Blanton, S. H. *et al. Genomics* **11**, 857–869 (1991).
28. Yang-Feng, T. L., Opdenakker, G., Voldkaert, G. & Francke, U. *Cytogenet. cell. Genet.* **40**, 784 (1985).
29. Polymeropoulos, M. H., Rath, D. S., Zhao, H. & Merril, C. R. *Nucleic Acids Res.* **19**, 969 (1991).
30. Aram, H. & Fataoureh, V. *Cutis* **14**, 215–218 (1974).
31. Cerimele, D. *et al. Hum. Genet.* **62**, 25–30 (1982).
32. Samantray, S. K., Samantray, S., Johnson, S. C. & Bhaktaviziam, A. *Aust. N.Z. J. Med.* **7**, 309–311 (1977).
33. Sparkes, R. S. *et al. Genomics* **1**, 138–144 (1987).

SUPPLEMENTARY INFORMATION. Requests should be addressed to the London editorial office of *Nature*.

ACKNOWLEDGEMENTS. This work was supported by Genentech Inc., by the Human Science Foundation, Japan, and by a grant from the NIH (J.W.). We thank K. Nishizawa and J. Tomfohrde for technical assistance; the many individuals from families affected with Werner's Syndrome; the Japanese-American blood donors; the oligonucleotides synthesis group at Genentech; S. Aotsuka, Y. Horiuchi, T. Tada, K. Okumura, K. Nishioka, Y. Mizushima, D. Botstein and R. White for encouragement; and the physicians and scientists in Japan who assisted in the collection of samples from affected families.

Transmission of chromosomal instability after plutonium α -particle irradiation

M. A. Kadhim, D. A. Macdonald, D. T. Goodhead, S. A. Lorimore, S. J. Marsden & E. G. Wright*

MRC Radiobiology Unit, Chilton, Didcot, Oxfordshire OX11 0RD, UK

WHEN investigating the biological effects of ionizing radiation on the haemopoietic system, a confounding problem lies in possible differences between the biological effects of sparsely ionizing, low linear energy transfer radiation such as X-, β - or γ -rays, and densely ionizing, high linear energy transfer radiation such as α -particles. To address this problem we have developed novel techniques for studying haemopoietic cells irradiated with environmentally relevant doses of α -particles from a plutonium-238 source. Using a clonogenic culture system, cytogenetic aberrations in individual colonies of haemopoietic cells derived from irradiated stem cells have been studied. Exposure to α -particles (but not X-rays) produced a high frequency of non-clonal aberrations in the clonal descendants, compatible with α -emitters inducing lesions in stem cells that result in the transmission of chromosomal instability to their progeny. Such unexpected instability may have important implications for radiation leukaemogenesis.

In mammals, the maintenance of cells in the peripheral blood is achieved by the proliferation and differentiation of precursor cells, all derived from a small, self-maintaining population of multipotential stem cells. Although it is evident that the stem cell compartment is heterogeneous with respect to a number of biological properties¹, the transplantation assay that detects murine spleen colony-forming units (CFU-S) is a useful measure of stem cells^{1,2}. We have demonstrated that a clonogenic assay for a cell (CFU-A) that has properties indistinguishable from those of CFU-S, provides a useful quantitative *in vitro* assay for these cells³. Having now developed a technique for irradiating haemopoietic cell suspensions in thin layers with Pu-238 α -particles incident as a parallel beam of energy 3.3 MeV and linear energy transfer (LET) 121 keV μm^{-1} , that is, near to the expected maximum biological effectiveness⁴, and established a technique for obtaining chromosome preparations from individual colonies, we have been able to do a karyotypic analysis of the effects of α -particle irradiation on murine stem cells.

Preliminary studies indicated that after exposure of normal murine bone marrow cells to 0.25, 0.5 or 1.0 Gray (Gy) Pu-238 α -particle irradiation, the proportions of surviving stem cells assayed as Day 12 CFU-S were roughly 65, 42 or 18% ($D_0 = 0.58$ Gy) (S.A.L., D.T.G. and E.G.W., manuscript in preparation) and were similar for CFU-A. The particle fluence was $5.15 \times 10^{10} \text{ m}^{-2} \text{ Gy}^{-1}$ so these doses correspond to a mean number of α -particles passing through a 7- μm diameter cell of 0.5, 1.0 or 2.0. Hence, many of the surviving cells were those which, by chance, did not receive a particle. A particular feature of irradiation of tissues by environmentally important α -particle-emitters is that it is entirely concentrated into a small number of separate, densely ionizing tracks of very limited ranges (35–90 μm). At low doses, any individual cell is likely to receive no dose or, if it happens to be in the path of a track, to receive a substantial dose of radiation (~ 0.5 Gy). For such α -emitters, therefore, the problem of whether low doses might be leukaemogenic reduces essentially to assessing the effectiveness of a single track, or a small number of tracks, in producing appropriate damage in the relevant target cell. Thus, for a stem cell with the properties of CFU-S that has an estimated diameter of 7 μm (ref. 5, 6), the doses we have used are directly relevant to the low-dose problem. Our survival data indicate that the probability of a stem cell (measured as CFU-S) surviving the passage of a single α -particle is only about 10% on the basis of simple biophysical considerations, assuming 7- μm diameter cells (S.A.L. *et al.*, manuscript in preparation). These surviving cells may carry viable genetic damage resulting from insult that is very much greater, at both the cell and chromatin level, than would ever be received from low doses of low-LET radiation⁷.

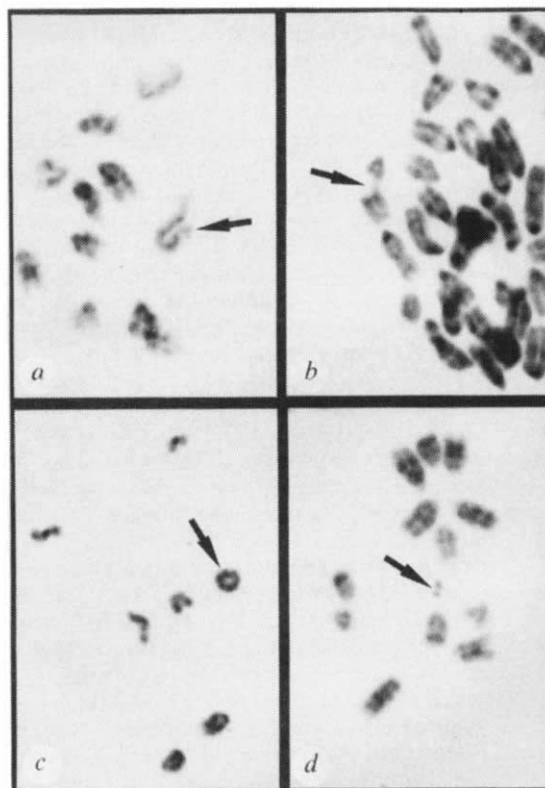


FIG. 1 Representative parts of metaphase spreads in colonies derived from CFU-A in marrow suspensions exposed to α -particle irradiation. *a, b*, Arrows single chromatid aberrations (respectively, a chromatid intrachange and isochromatid deletion). *c, d*, Arrows, single chromosome aberrations (respectively, a ring and a double minute, which may be secondarily derived from previous chromatid changes).

* To whom correspondence should be addressed.

TABLE 1 Cytogenetic aberrations in colonies derived from stem cells in marrow suspensions exposed to α -particle- or X-irradiation

Radiation exposure (Gy)	Colonies with aberrations	Metaphases with aberrations	Aberrations/metaphase				Total aberrations	Number of different aberrations	
			1	2	3	>3		chromosome	chromatid
0	7/59	7/432	7	0	0	0	7	0	7
α Particles									
0.25	2/5	8/29	7	0	1	0	10	3	7
0.50	6/12	19/92	13	3	2	0	25	8	14
1.00	4/10	26/107	14	7	4	1	44	4	34
X-rays									
3.0	2/86	2/409	2	0	0	0	2	0	2
		16/409	2	0	0	0	16	2	0
			(clonal)						

Bone marrow cells from male CBA/H mice were irradiated with α particles at 0.2–0.8 Gy min⁻¹ using a recently constructed variable dose-rate source containing a 20-mm-diameter disc of plutonium-238 (ref. 16). Cells were X-irradiated at 0.75 Gy min⁻¹ using a Siemens Stabilipan X-ray machine operating at 250 kV constant potential and 14 mA, giving a half value layer of 1.2 mm of copper. Immediately after irradiation the cells were washed, resuspended and the CFU-A assay³ was used to obtain colonies of cells derived from members of the haemopoietic stem cell compartment. Cells were plated in 45-mm Petri dishes containing 2 ml modified alpha Eagles medium supplemented with 20% horse serum, 0.3% agar, antibiotics, glutamine and sources of colony-stimulating activities as described previously³. For quantitative studies, triplicate cultures were incubated at 37 °C in a fully humidified atmosphere of 10% CO₂, 5% O₂ and 85% N₂ for 11 days. Cytogenetic analyses were done using a previously reported method for karyotyping haemopoietic colonies¹⁷ and in our experiments, only colonies that met the size criterion for being scored as CFU-A-derived³ were selected for study. Briefly, metaphases in developing day 7 colonies (containing 10³–10⁴ cells, that is, some 10–13 cell divisions from initiation of clonal proliferation) were arrested by adding colcemid to the dishes. Individual colonies were transferred in 10- μ l droplets of 0.5% KCl onto poly-L-lysine coated microscope slides and hypotonic treatment of the cells was achieved by inverting the slide to prevent attachment and allowing the cells to swell in a hanging droplet. After 25 min in a humidified incubator at 37 °C, the slide was turned upright and the cells allowed to attach to the coated surface of the slide. Air-dried slides were fixed and G-banded using standard methods¹⁸.

Metaphase preparations of colonies produced by CFU-A surviving α -particle irradiation revealed that 40–60% had karyotypic abnormalities (Table 1) and chromatid aberrations appeared at a greater frequency than chromosome aberrations. In an individual colony, not all cells exhibited abnormalities and up to 50% of scorable metaphases may carry single or multiple, nonidentical aberrations; that is, no aberrations were clonal. Examples of aberrations found in colonies are shown in Figs 1 and 2. The results of experiments in which we exposed the cells to 3 Gy X-rays, a dose that reduced the survival of CFU-A to about 5%, were markedly different to the results of our Pu-238 experiments. Of 86 CFU-A-derived colonies studied in detail, only two had chromosome abnormalities. In both cases they were clonal aberrations; that is, present in all scorable metaphases from the colony. Two cells from different colonies had single chromatid aberrations; an incidence comparable to the background frequency in control colonies.

Our results demonstrate a significant and important difference in the effects of the two types of irradiation. It is evident that the abnormalities observed in the colonies derived from cells surviving α -particle irradiation are present at high frequencies, comparable to those of exchange chromosome aberrations in α -irradiated human blood lymphocytes⁸. But, most strikingly, the nonclonality and variable number of cells in a colony exhibiting chromatid aberrations is consistent with them arising *de novo* in cells derived from a clonogenic cell that survived the passage of one or more radiation tracks before the initiation of clonal proliferation. For most biological endpoints the relative biological effectiveness (RBE) of slow α -particles, relative to low-LET radiations, is 3–50 (refs 4, 9, 10) and the currently recommended weighting factor for radiological protection is 20 (ref. 11). Our experiments provide evidence for a unique effect of α -particles, suggesting an effective RBE approaching infinity as was also suggested for sister chromatid exchanges in resting human lymphocytes¹².

The pattern of Pu-238 α -particle-induced karyotypic abnormalities we have demonstrated suggests that the exposed, surviving stem cells transmit to their daughter cells some chromosomal instability that may result in one or more visible cytogenetic aberrations many cell cycles later. We have no evidence to suggest that particular chromosomes are consistently involved, or that these aberrations arise at 'fragile sites' and it is possible

to hypothesize that the instability could, on occasions, disrupt a region of the genome involved in leukaemic transformation. As stem cells have the property of self-renewal, such a change could arise in a daughter stem cell and thus represent an apparent 'initiating lesion'. The actual initiating lesion is, of course, the radiation-induced instability some generations earlier.

There is much concern about clusters of leukaemia

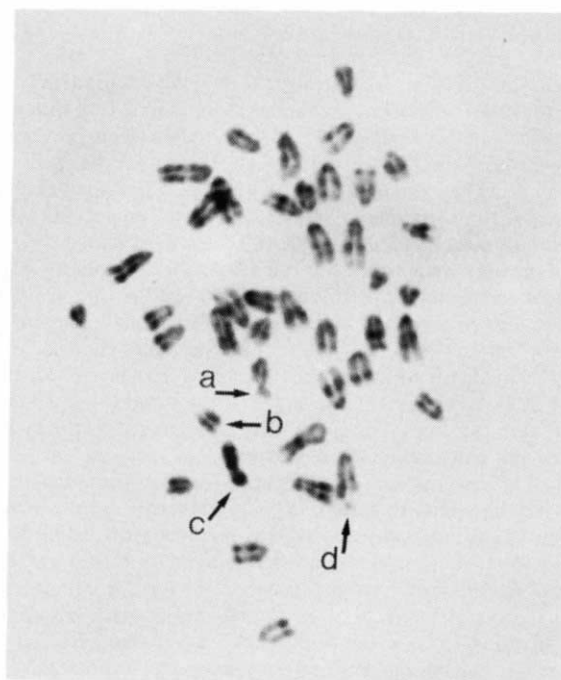


FIG. 2 An example of a metaphase spread exhibiting multiple aberrations in a colony derived from a CFU-A in a marrow suspension exposed to α -particle irradiation. Arrowed are: a, chromatid break; b, acentric chromosome fragment; c, possible derivative chromosome with either constriction or median centromere (probably translocation-derived); d, single minute chromosome.

epidemiologically linked to nuclear sites^{13,14}. Estimates of radiation dose in these cases are based on data from environmental monitoring used in conjunction with complex environmental and biological pathway models. Estimates of leukaemia risk are based conventionally on epidemiological data for low-LET radiations and enhanced effectiveness of high-LET radiations from experimental RBE measurements¹⁵. These estimates are not consistent with a causal connection between the doses and the leukaemias^{13,14}. But if as we believe from theoretical biophysical considerations⁷ and some experimental support¹², that there may be classes of unique, initial radiogenic damage

induced only by high-LET radiations, then the RBE for such damage would be effectively infinite. Leukaemias arising from such a situation may not have been identified as radiogenic from human epidemiological data (which is based predominantly on considering low-LET radiation) and our findings may then have considerable relevance to the problem of low-dose radiation exposure from artificial or natural α -emitters. Investigating the longer-term consequences of the demonstrated chromosomal instability and identifying the molecular basis of the α -particle-induced lesion are significant challenges for future studies. □

Received 24 September; accepted 25 November 1991.

- Visser, J. W. M. & Van Bekkum, D. W. *Expl Hematol.* **18**, 248–256 (1990).
- Till, J. E. & McCulloch, E. A. *Radiation Res.* **14**, 213–222 (1961).
- Lorimore, S. A., Pragnell, I. B., Eckmann, L. & Wright, E. G. *Leuk. Res.* **14**, 481–489 (1990).
- Goodhead, D. T. in *Dosimetry of Ionizing Radiations* Vol. 2 (eds Kase, K. R., Bjarngaard, B. E. & Attix, F. H.) 1–89 (Academic, New York, 1987).
- Visser, J., van den Engh, G., Williams, N. & Mulder, D. in *Experimental Hematology Today* (eds Baum, S. J. & Ledney, G. D.) 21–27 (Springer, New York, 1977).
- van den Engh, G., Visser, J. & Trask, B. in *Experimental Hematology Today* (eds Baum, S. J. & Ledney, G. D.) 19–26 (Springer, New York, 1979).
- Goodhead, D. T. *Int. J. radiat. Biol.* **56**, 623–634 (1989).
- Lloyd, D. C. & Edwards, A. A. in *Radiation-Induced Chromosome Damage in Man* (eds Ishihara, T. & Sasaki, M. S.) 23–49 (Liss, New York, 1983).
- Sinclair, W. K. *Radiat. protec. Dosim.* **13**, 319–329 (1985).
- NCRP *The Relative Biological Effectiveness of Radiations of Different Quality*, NCRP Report 106 (National Council on Radiation Protection and Measurement, Bethesda, 1990).

- ICRP 1990 *Recommendations of the International Commission on Radiological Protection*, ICRP Publication 60 (Pergamon, Oxford, 1991).
- Aghamohammadi, S. Z., Goodhead, D. T. & Savage, J. R. K. *Int. J. Radiat. Biol.* **53**, 909–915 (1988).
- Black, D. (Chairman) *Investigation of Possible Increased Incidence of Cancer in West Cumbria: Report of the Independent Advisory Group* (HMSO, London, 1984).
- Bobrow, M. (Chairman) *Committee on Medical Aspects of Radiation in the Environment (COMARE): Third Report* (HMSO, London, 1989).
- UNSCEAR *United Nations Scientific Committee on the Effects of Atomic Radiation: Report to the General Assembly* (United Nations, New York, 1988).
- Goodhead, D. T., Bance, D. A., Stretch, A. & Wilkinson, R. E. *Int. J. Radiat. Biol.* **59**, 195–210 (1991).
- Amenomori, T. et al. *Int. J. Cell Cloning* **3**, 133–142 (1985).
- Piper, J. & Breckon, G. *Cytogenet. Cell Genet.* **50**, 111–115 (1989).

ACKNOWLEDGEMENTS. We thank R. Wilkinson, P. Adams and D. Badger for technical assistance, and J. Savage and E. Evans for discussion. This work was supported, in part, by the United Kingdom Coordinating Committee for Cancer Research.

Crystal structure of a dUTPase

Eila S. Cedergren-Zeppezauer*, Gunilla Larsson†, Per Olof Nyman†, Zbigniew Dauter‡ & Keith S. Wilson‡

* Department of Zoological Cell Biology, Wenner-Gren Institute, University of Stockholm, S-10691 Stockholm, Sweden

† Department of Biochemistry, Chemical Centre, University of Lund, S-22100 Lund, Sweden

‡ European Molecular Biology Laboratory (EMBL), c/o DESY, Notkestrasse 85, D-2000 Hamburg 52, Germany

THE enzyme dUTPase catalyses the hydrolysis of dUTP¹ and maintains a low intracellular concentration of dUTP so that uracil cannot be incorporated into DNA². dUTPase from *Escherichia coli* is strictly specific for its dUTP substrate,³ the active site discriminating between nucleotides with respect to the sugar moiety as well as the pyrimidine base. Here we report the three-dimensional structure of *E. coli* dUTPase determined by X-ray crystallography at a resolution of 1.9 Å. The enzyme is a symmetrical trimer, and of the 152 amino acid residues in the subunit, the first 136 are visible in the crystal structure. The tertiary structure resembles a jelly-roll fold and does not show the 'classical' nucleotide-binding domain. In the quaternary structure there is a complex interaction between the subunits that may be important in catalysis. This possibility is supported by the location of conserved elements in the sequence.

The dUTPase enzyme (EC 3.6.1.23) was obtained using an overproducing genetic construct⁴. Crystals were grown at room temperature in a mixture of polyethyleneglycol PEG8000, 50 μ M MgCl₂, 450 μ M pyrophosphate and succinate buffer at pH 4.2 (ref. 5). At this pH, activity measurements are difficult, but dissolved crystals show enzyme activity from pH 5. These crystals diffract X-rays to beyond 1.7 Å resolution. There is one subunit per asymmetric unit (data collection summarized in Table 3). Details of the crystallographic work will be published elsewhere.

dUTPase is packed as a trimer around the threefold axis of the crystal, with large solvent channels between the enzyme molecules. Chromatography confirms that the molecular weight of dUTPase in solution corresponds to a trimer (G.L., manuscript in preparation), rather than to a tetramer as reported

TABLE 1 X-ray data statistics

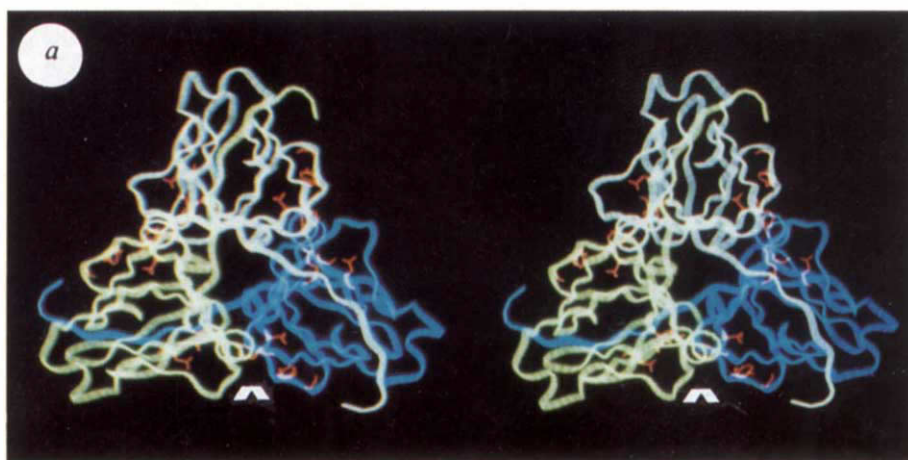
	Resolution (Å)	λ (Å)	R-merge (%)	Unique reflections	Metal sites per subunit
Native E(Mg-PP _i) data: 2 crystals	1.9	1.009	6.9	13,640 96% complete	None found
Derivatives:					
Hg, 1 crystal	2.0	0.995	5.9	11,730	1
Pt, 1 crystal	2.1	1.050	5.7	9,910	1

The space group is *R*3 with cell dimensions $a=b=86.6$ Å and $c=62.3$ Å, and the crystals contain one subunit per asymmetric unit. Intensity data were collected using monochromatic synchrotron radiation at the EMBL outstation at DESY, Hamburg. The Hendrix-Lentfer imaging plate scanner, constructed in-house at EMBL, was used as detector. Data were processed with a modified version of the MOSFLM package. R-merge is defined as $\sum |I_i - \langle I \rangle| / \sum \langle I \rangle$, where I_i is an individual intensity measurement and $\langle I \rangle$ is the mean intensity for this reflection. Derivative data were collected at wavelengths chosen to maximize the anomalous signal. For the mercury derivative, a native crystal, E(Mg-PP_i), was soaked in 75 μ M ethyl mercury phosphate for 4 h. The second platinum derivative was similarly prepared using 1.2 mM K₂PtCl₄ for 7 h. The mercury compound reacted with the only cysteine residue present in the sequence, whereas platinum binds to a methionine side chain. Both isomorphous and anomalous components were used for the phase determination to 2.2 Å resolution, giving an overall figure of merit of 51%, which was increased to 78% by solvent flattening¹⁶. This gave a high-quality map with clear continuity in the electron density. The model was built using the program FRODO¹⁷ on an Evans and Sutherland PS330 graphics system. The initial model was refined by a cycle of simulated annealing using XPLO¹⁸, and subsequently by restrained least-squares minimization using PROLSQ¹⁹. The final R factor for 136 residues and 189 water molecules is 14.5% when all reflections between 1.9 and 8.0 Å are used. The geometry of the final model is in good agreement with the target values for the stereochemical parameters: the deviation in bond lengths was 0.010 Å and in torsion angles 2.5°. The average temperature factor for the protein atoms was 20.9 and for the solvent molecules 38.1.

earlier⁶. A trimeric subunit arrangement has also been found for a mammalian dUTPase⁷. The trimer has a wedge-shaped appearance when viewed perpendicularly to the 3-fold axis and a triangular face, typical for a trimer, when viewed along the 3-fold axis (Fig. 1a). The largest distance across a triangular face is about 60 Å and the length along the axis is ~45 Å. The subunits associate through interactions between twisted β -sheets, thus burying three hydrophobic surfaces about the 3-fold axis. The arrangement of monomers resembles that of the trimeric tumour necrosis factor^{8,9}, but the details of relative orientation of strands along the axis differ.

Figure 2a and b shows a topology diagram and a ribbon representation of the subunit which consists of a polypeptide

FIG. 1 *a*, Stereo view along the 3-fold axis of a ribbon diagram of trimeric dUTPase. The three subunits are shown in different colours. Side chains are only shown for those conserved residues that are accessible to solvent in the trimer. The depression on the surface proposed to contain one active site of dUTPase is marked by the white arrows. *b*, Electron-density map of the wedge-shaped subunit viewed at an angle to the 3-fold axis, with the model of 136 residues superimposed. Although hydrophobic residues pack closely within the jelly-roll, this core at the centre of the subunit is surrounded by small hydrophobic cavities. The synthesis is contoured at a level of 1σ above the average density.

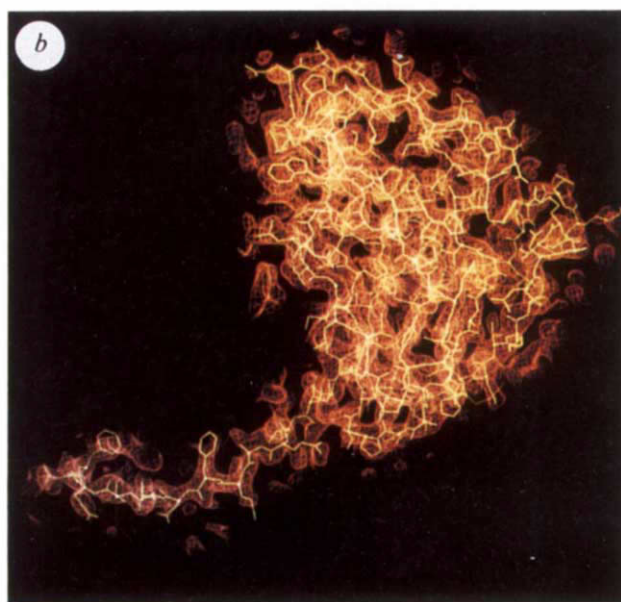


chain of 152 amino acid residues; the first 136 of these are visible in the crystal structure. The remaining 16 residues up to the carboxy terminus contain several conserved side chains, but cannot be seen in the X-ray structure. The subunit is made up of about 40%, mainly antiparallel, β -pleated sheet and has a single short α -helix. The major tertiary structural feature is a sandwich made by two β -pleated sheets, 10–13 Å apart. The volume between them composes the hydrophobic core of the subunit. Two antiparallel β -strands, which are separate from the sandwich, together with one short helix and a long loop, complete the fold of the subunit. The structure resembles an 8-stranded jelly-roll composed of parts from two subunits and distorted by the direction of strand 1 (the amino terminus) being opposite to that of a regular jelly-roll¹⁰.

The first two methionines in the dUTPase sequence, Met 1 and Met 2, are poorly defined in the electron density map. From residue 3 to 136 (Fig. 1*b*) the initial map was easy to interpret. Density is clearly visible for the side-chain ring of Phe 136, partly embedded among side chains of a symmetry-related subunit in the trimer (Fig. 3), but electron density abruptly stops after this residue. The remaining 16 amino acids of dUTPase are nevertheless present in the protein crystals (G.L., manuscript in preparation). Phe 136 lies on the surface of the trimer exposed to solvent channels in the crystal, and the 'missing' residues are presumed to be disordered.

The interface between adjacent subunits in the trimer is made up by residues from a number of different parts of the sequence. (1) A hydrophobic interaction around the 3-fold axis at the centre of the trimer involves residues 47–58, β -strand 3 in the topology diagram, and 96–103, β -strand 6. Along the central axis, side chains of methionines, leucines, isoleucines, valines and threonines are closely packed. (2) The residues 20–30 of the loop region in one subunit form a contact to the bend 90–94 in another subunit (Fig. 2*b*). (3) A 35 Å-long carboxy terminal arm (residues 127–136) extends from the core of one subunit and makes extensive contacts along the surface of the neighbouring subunit (Fig. 3*a*). It constitutes strand 8 in the jelly-roll model (Fig. 2*a*), and forms the fifth strand in one of the β -sheets. A similar situation is encountered for the β -subunit of heavy riboflavin synthase¹¹, where the amino terminus forms the fifth strand in the β -sheet of a neighbouring subunit in a pentamer.

The enzyme requires divalent metal ions such as Mg^{2+} for activity, and the activation constant for Mg^{2+} ion and the inhibition constant for the cleavage product, pyrophosphate, have been reported for *E. coli* dUTPase¹². These ions were included in the crystallization medium with the intention of forming a complex allowing the identification of the active site. Mg^{2+} and pyrophosphate cannot be seen in the electron density maps.



Given the high resolution and low *R*-factor (Table 1), this clearly shows that they are not bound in an ordered manner in our crystals, suggesting that the constants estimated differ from binding constants, at least in the crystal. This might be coupled to the fact that the carboxy terminus is flexible. Under certain conditions, the disordered residues may partake in the formation of a binding site for the nucleoside triphosphate that is made up from elements from more than one subunit.

Sequence data on dUTPases have mainly been derived from the DNA of the corresponding gene. The sequence of the *E. coli* enzyme¹³ has been compared with that of three different herpesviruses¹⁴ (herpes simplex, herpes zoster varicellous, Epstein-Barr virus). Five regions, scattered along the sequence, are highly conserved (Fig. 2*b* and *c*), indicating that those parts are important for the catalytic function and/or stability. An inspection of the three-dimensional model of *E. coli* dUTPase reveals that some of the conserved residues are located in bends and loops that are in van der Waals contact. The sequence of residues 137–152 (not visible in the electron density maps) is particularly rich in glycines and shows significant homology not only to viral dUTPases but also to phosphate-binding sequences in other nucleotide-binding proteins¹⁵. But, the dUTPase subunit does not show the 'classical' Rossmann fold and thus represents a novel polypeptide fold for nucleotide binding.

Three shallow depressions are observed on the surface of the trimer. They can be seen at the interface between two of the

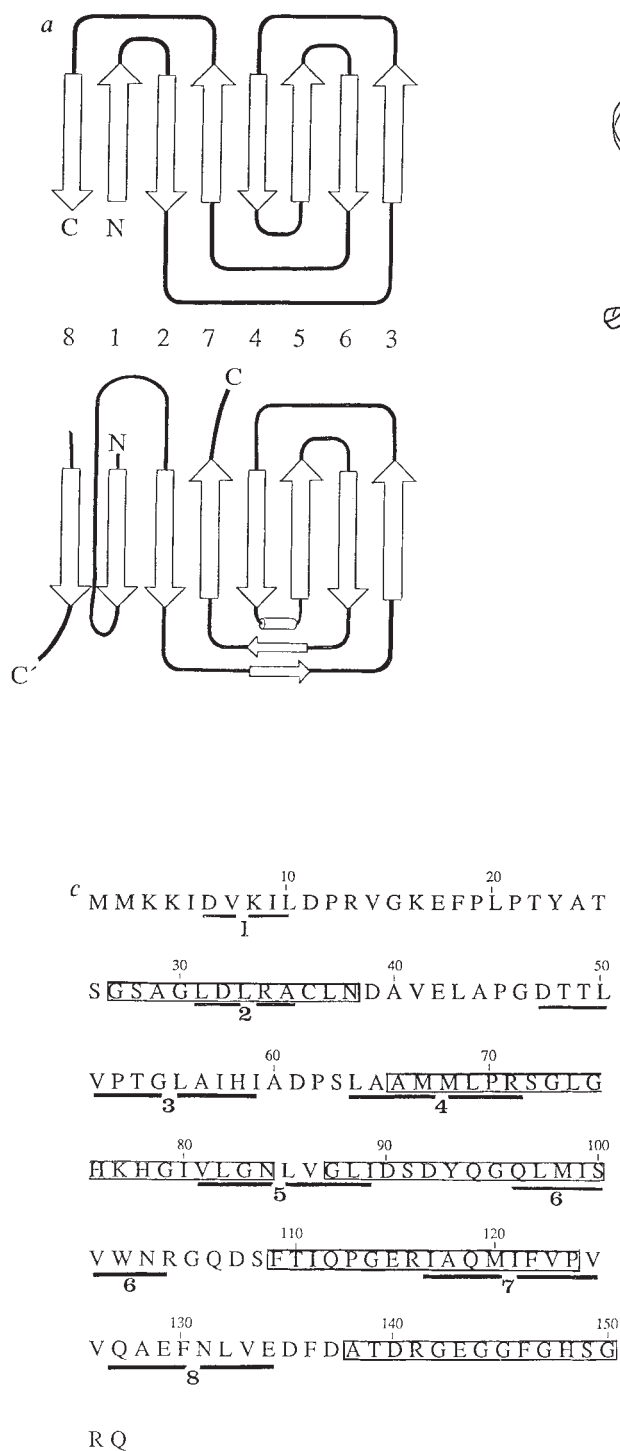


FIG. 2 *a*, Topology diagram of an idealized jelly-roll fold¹⁰ with strand numbers indicated (top) compared with that of the dUTPase subunit. N and C, amino and carboxy termini; C' indicates the carboxy terminus of an adjacent subunit. *b*, Schematic drawing of the subunit. Arrows represent β -strands. The first and last amino-acid residue of each strand within the jelly-roll barrel are labelled. Strand 8' is from a neighbouring subunit. Shaded parts of the structure correspond to conserved sequences. *c*, The amino-acid sequence of dUTPase from *E. coli*. Residues 137–152 are not visible in the electron density map. To confirm the existence of these last 16 residues, peptides generated by tryptic hydrolysis of dissolved crystals were isolated and subjected to Edman degradation. Sequence alignments of dUTPases from *E. coli* and various viruses indicate the presence of homologous regions. Strongly conserved regions are boxed, and correspond to the shaded parts in *b*. Underlined sequences, labelled with bold numbers, show β -strand numbering according to the topology diagram in *a* and the same convention holds for the strand numbering in *b*.

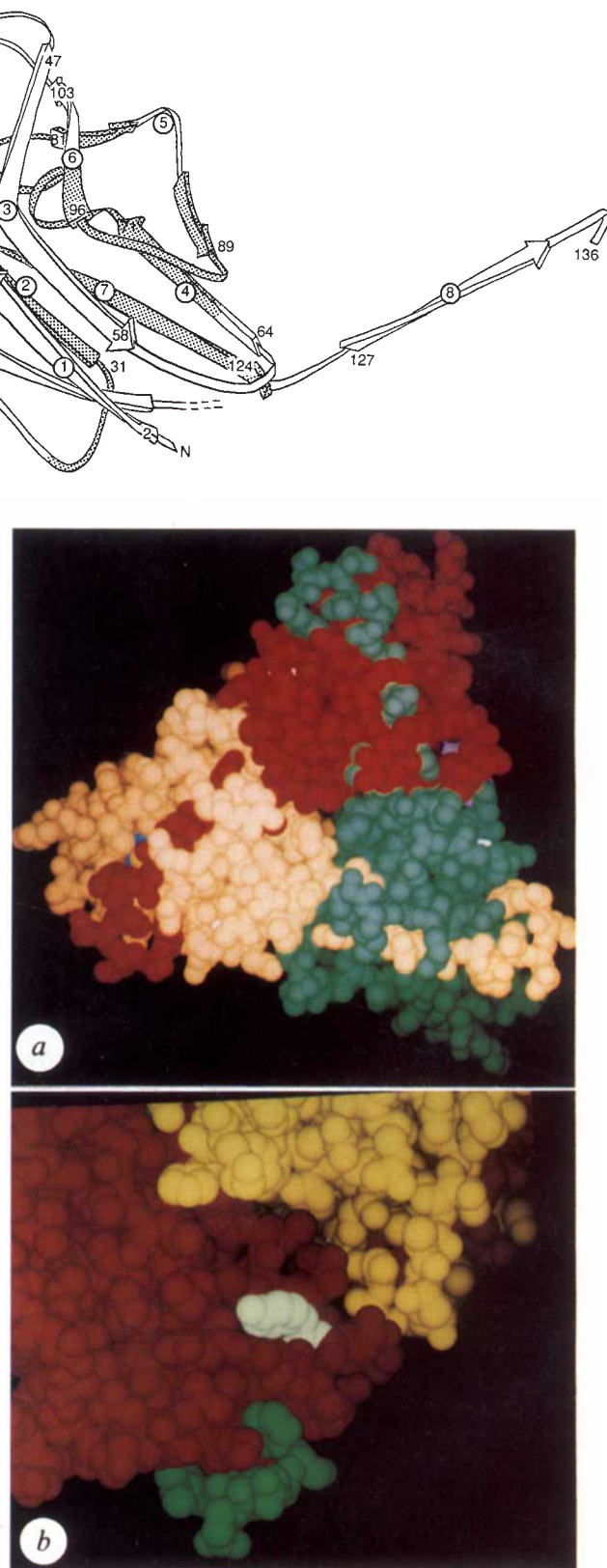


FIG. 3 Space-filling model of the dUTPase trimer viewed along the 3-fold axis with subunits colour-coded, showing carboxy terminal 'arms' running across the surface of a neighbouring subunit: for example, a green 'arm' passing the red subunit and thereafter becoming disordered. *b*, The molecule rotated to show a close up of the intersubunit contacts at one of the three surface cavities. Evolutionarily conserved loops meet to form the depression comprising the bend 90–94 (red subunit, with the strictly conserved Tyr 93 coloured white) and the loop 20–30 (yellow subunit). Towards the depression a carboxy terminal 'arm' approaches (green subunit).

subunits when viewed down the 3-fold axis as in Fig. 1a. The side chains of eight conserved amino acids (residues Gly 30, Asp 32, Ser 72, Ile 89, Asp 90, Tyr 93, Gly 95 and Gln 119), which have at least 10 Å² surface area accessible to solvent in the trimer, are seen to lie in and around these depressions. Moreover, the end of the extended arm (Phe 136) of the third subunit points towards this region (Fig. 3b). The conserved invisible carboxy terminus (Fig. 2c) can thus be imagined as interacting with the shallow cavity region. Although the active site cannot be identified with certainty, the information accumulated so far indicates that it could lie in that region. □

Received 5 September; accepted 19 November 1991.

- Kornberg, A. & Baker, T. *DNA Replication* (Freeman, New York, 1991).
- Tye, B. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **74**, 154-157 (1987).
- Bertani, L. E., Häggmark, A. & Reichard, P. *J. J. Biol. Chem.* **238**, 3407-3413 (1963).
- Hoffmann, I. et al. *Eur. J. Biochem.* **164**, 45-51 (1987).
- Cedergren-Zeppezauer, E. S. et al. *Proteins: Struct. Funct. Genet.* **4**, 71-75 (1988).
- Shlomai, J. & Kornberg, A. *J. Biol. Chem.* **253**, 3305-3312 (1978).
- Hokari, S. & Sakagishi, Y. *Arch. biochem. Biophys.* **253**, 350-356 (1987).
- Jones, E. Y., Stuart, D. I. & Walker, N. P. C. *Nature* **338**, 225-228 (1989).
- Eck, M. J. & Sprang, S. R. *J. Biol. Chem.* **264**, 17595-17605 (1989).
- Branden, C.-I. & Tooze, J. in *Introduction to Protein Structure* 70-74 (Garland, New York, 1991).
- Ladenstein, R. et al. *J. molec. Biol.* **203**, 1045-1070 (1988).
- Hoffmann, I. thesis, Univ. Saarbrücken, Germany (1988).
- Lundberg, L. G., Thoreson, H.-O., Karlström, O. & Nyman, P. O. *EMBO J.* **2**, 967-971 (1983).
- McGeoch, D. J. *Nucleic Acids Res.* **18**, 4105-4110 (1990).
- Möller, W. & Amos, R. *FEBS Lett.* **186**, 1-7 (1985).
- Wang, B. C. *Meth. Enzym.* **115**, 90-112 (1985).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268-272 (1978).
- Brunker, A. T. *J. molec. Biol.* **203**, 803-816 (1988).
- Hendrickson, W. A. & Konner, J. H. in *Biomolecular Structure, Conformation, Function and Evolution* Vol. 1 (ed. Srinivasan, R.) 43-57 (Pergamon, Oxford, 1980).

ACKNOWLEDGEMENTS. We thank A. M. Rosengren for enzyme purification facilities; C.-I. Brändén for suggestions and reading the manuscript; B. Furugren, S. Al-Karadaghi and T. Astlind for artwork; and M. Dauter for technical assistance. This work was supported by the Swedish Natural Science Research Council, Kungliga Fysiografiska Sällskapet i Lund, Swedish Cancer Society, and Wallenbergstiftelsens Jubileumsfond (G.L. and E.C.-Z.).

Renaturation of DNA catalysed by yeast DNA repair and recombination protein RAD10

Patrick Sung, Louise Prakash* & Satya Prakash†

Department of Biology, University of Rochester, River Campus Station, Rochester, New York 14627, USA

* Department of Biophysics, University of Rochester School of Medicine, 601 Elmwood Avenue, Rochester, New York 14642 USA

THE RAD10 gene of *Saccharomyces cerevisiae* is required for the incision step of excision repair of ultraviolet-damaged DNA (refs 1, 2), and it functions in mitotic recombination³. RAD10 has homology to the human excision repair gene ERCC-1 (ref. 4). Here we describe the purification of the protein encoded by RAD10 and show that it is a DNA-binding protein with a strong preference for single-stranded DNA. We also show that RAD10 promotes the renaturation of complementary DNA strands.

RAD10 protein, overproduced in yeast strain CMY135 carrying plasmid pSUC8 (Fig. 1a), was purified to near homogeneity (Fig. 1b). The relative molecular mass as estimated by SDS-PAGE of purified RAD10 protein is 25,000 (*M*_r 25 K), agreeing well with the expected size of 24.3 K⁵. The DNA-binding properties of RAD10 were characterized using a nitrocellulose filter binding assay. The percentage of single-stranded (ss) DNA retained on alkali-treated filters⁶ was proportional to the amount of RAD10, reaching a maximum of just over 80% of input DNA at 75 ng RAD10 (Fig. 2), corresponding to a protein : DNA molar ratio of about 145 or one RAD10 monomer per 50 nucleotides. The binding of ssDNA is unaffected by KCl concentrations up

to 50 mM, with 75 and 50% of the binding activity remaining at 100 and 150 mM KCl, respectively. By contrast, RAD10 showed little binding to double-stranded (ds) DNA (Fig. 2). A strong preference for binding ssDNA was also observed in mobility shift assays in agarose gels (data not shown). Ultraviolet irradiation (1,000 J m⁻²) of DNA did not stimulate binding of RAD10 to ss- or dsDNA (Fig. 2). In mobility-shift assays on nondenaturing polyacrylamide gels, there was also no preferential binding of RAD10 to a ³²P-labelled ultraviolet-irradiated 70-base-pair (bp) DNA fragment containing a stretch of thymine residues to aid the formation of thymine dimers (data not shown). The middle portion of RAD10 contains many tyrosine residues and is basic⁵. This region of RAD10 could be involved in ssDNA binding by intercalation of tyrosine residues between the bases in DNA^{7,8} and through ionic interactions. The ssDNA-binding activity coeluted with RAD10 during the last step of purification on Sephadex G75 (data not shown). This and the high purity (>98%) of RAD10 strongly suggest that the ssDNA-binding activity is intrinsic to RAD10.

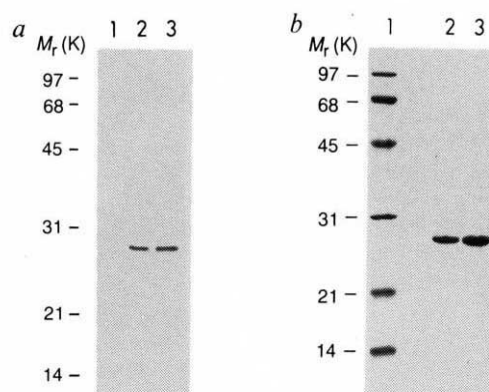


FIG. 1 Overproduction and purification of RAD10. a, The nitrocellulose blot of a 12% denaturing polyacrylamide gel was probed with affinity-purified anti-RAD10 antibodies. Lane 1, extract from yeast strain CMY135 carrying the vector pSCW231 (2 μ plasmid replication origin, *ADCI*) that lacks the *RAD10* gene; lane 2, extract from CMY135 strain carrying the *RAD10* overproducing plasmid pSUC8 (2 μ , *ADCI-RAD10*); lane 3, 20 ng purified RAD10 protein. b, Purified RAD10 protein was electrophoresed on a 12% denaturing polyacrylamide gel and stained with Coomassie blue. Lane 1, *M*_r markers; lane 2, 2 μ g RAD10 protein; lane 3, 4 μ g RAD10 protein.

METHODS. For raising polyclonal antibodies against the RAD10 protein, plasmid pPS7 was used to produce the ρ -RAD10 fusion protein in *E. coli*. In pPS7, the RAD10 protein coding region missing the first 6 codons is fused in frame to the first 116 codons of the ρ gene which is under the control of the λ P₁ promoter. Plasmid pSUC8 was used for overproducing RAD10 in yeast. In this plasmid, the *RAD10* gene from position -37 to 156 nucleotides downstream of the TGA termination codon⁵ is fused to the *S. cerevisiae* highly expressed constitutive *ADC1* promoter. Breakage of yeast cells by the French press was as described previously²⁴. Unless stated otherwise, all chromatographic columns were equilibrated in buffer A (20 mM KH₂PO₄ (pH 7.0) containing 10% glycerol, 1 mM DTT and 0.5 mM EDTA). Extract from 130 g CMY135(pSUC8) was treated with 0.31 g ml⁻¹ ammonium sulphate, and the pellet was dissolved in 80 ml buffer A and dialysed overnight against 3 l of buffer A. The dialysate was passed through a column of DEAE Sephacel (5 $\times$ 10 cm) to remove particulate material and the bulk of unwanted proteins. The flow-through fraction from DEAE (fraction 1) was applied onto Bio-Rex 70 (2.5 $\times$ 12 cm), which was developed with a 600 ml gradient of 0-200 mM KCl. Fractions containing RAD10 (eluting between 130-145 mM KCl) were identified by immunoblotting, pooled (40 ml; fraction 2) and dialysed against buffer A. Proteins in the dialysate were fractionated on Mono S (HR 5/5; Pharmacia) with a 33 ml, 0-180 mM KCl gradient. RAD10 eluted from Mono S between 140-150 mM KCl, and the pool of which (4 ml; fraction 3) was concentrated to 0.8 ml with a Centricon-10 device (Amicon). Trace quantities of contaminating proteins in fraction 3 were removed by gel filtration on Sephadex G75 superfine (1 $\times$ 45 cm) equilibrated and eluted with buffer A containing 130 mM KCl. The overall yield was 400 μ g of apparently homogeneous RAD10 protein with a recovery of 20%.

† To whom correspondence should be addressed.

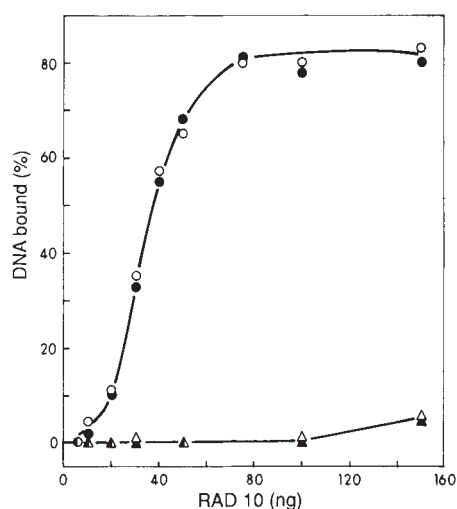


FIG. 2 RAD10 binds DNA. ^3H -labelled M13 viral ssDNA without (●) or with (○) ultraviolet treatment and ^3H -labelled M13 dsDNA (60% supercoiled and 40% nicked circular) without (▲) or with (△) ultraviolet treatment. METHODS. RAD10 protein (5–150 ng or 0.2–6 pmol) was incubated with DNA (50 ng or 153 pmol) for 15 min at 30 °C in 10 μl reaction buffer (20 mM potassium MES (pH 6.0) containing 60 $\mu\text{g ml}^{-1}$ BSA, 50 mM KCl, 10 mM MgCl_2 , 1 mM ATP, and 1 mM DTT). Omission of ATP and Mg^{2+} did not have any effect on RAD10 DNA binding. The reaction mixtures were diluted with 100 μl of ice-cold wash buffer (reaction buffer without BSA and ATP) and filtered dropwise through NaOH-treated⁶ HAWP nitrocellulose filters (Millipore) with the aid of a water suction pump. The filters were rinsed twice with 3 ml ice-cold wash buffer, dissolved in 10 ml Filtion X scintillation fluid (National Diagnostics), and assayed for radioactivity. For ultraviolet treatment, DNA (50 $\mu\text{g ml}^{-1}$) was irradiated in 10 μl droplets of H_2O at a fluence rate of 10 $\text{J m}^{-2} \text{s}^{-1}$ for 100 s, using 15 W germicidal bulbs emitting at 254 nm. Addition of 1% SDS to the reaction mixture before filtration through nitrocellulose abolished DNA binding by RAD10. In control experiments, both the ss- and dsDNAs were bound by the *E. coli* RecA protein with the expected efficiency²⁵.

Because of its role in genetic recombination³, we examined whether RAD10 promotes the renaturation of homologous single strands of DNA. Incubation of heat-denatured dsDNA with RAD10 formed a DNA species that remained at the top of 0.8% agarose gel (Fig. 3). Stripping RAD10 from DNA before electrophoresis by the addition of 1% SDS (Fig. 3) did not destabilize the high M_r DNA product, but boiling dissociated the DNA product into a form that had the same mobility as the heat-denatured DNA (Fig. 3, lane 6 versus lane 3). No formation of high M_r DNA occurred if the dsDNA was not heat-denatured (Fig. 3, lanes 1 and 2) or when noncomplementary M13 viral ssDNA was used (data not shown). Taken together, these observations suggest that RAD10 promotes the renaturation of complementary DNA strands. The RAD10-mediated renaturation reaction is not dependent on ATP and Mg^{2+} (Fig. 3, compare lane 5 with lanes 7–9). The DNA renaturation products were also examined by electron microscopy⁹ after removal of RAD10 by SDS treatment and phenol extraction. Large aggregates of DNA were formed on incubation of heat-denatured dsDNA with RAD10, but not in the control lacking RAD10 (data not shown).

RAD10-mediated renaturation of DNA was also examined by S1 nuclease digestion. There was a RAD10 concentration-dependent conversion of heat-denatured T7 DNA into an S1 resistant form (Fig. 4a), which reached a near maximal level of between 50–55% of input DNA (above the background value of 5% in the absence of RAD10) at 500 ng RAD10, or at a protein to DNA ratio of one RAD10 monomer per 60 nucleotides. Incubation of ^3H -labelled noncomplementary M13 viral ssDNA with RAD10 under identical conditions did not render the DNA S1 resistant, indicating a prerequisite for homology. The RAD10-catalysed DNA-renaturation reaction was rapid, as most of the S1-resistant product was formed within 2 min of incubation, and the kinetics of formation of S1-resistant DNA suggests that RAD10 acts in a stoichiometric fashion (Fig. 4b). The salt sensitivity of the DNA renaturation reaction mirrored that of ssDNA binding: it was unaffected by 50 mM KCl, and at 100 mM and 150 mM KCl, 70% and 50% of the renaturation activity remained, respectively.

Incubation of RAD10 protein (0.5–3 μg) in 10 μl reaction mixtures containing 100 ng M13mp18 viral ssDNA and 200 ng *Bgl*II-linearized M13mp18 dsDNA for 60 min did not result in joint molecule formation between the two DNA species, as examined by agarose gel electrophoresis, indicating that RAD10 does not possess strand-exchange activity. RAD10 does not hydrolyse ATP and other nucleotides in the presence or absence of DNA, and has no nuclease activity (data not shown).

In excision repair, the ssDNA-binding activity of RAD10 could be required after binding and unwinding of the damage site by the damage-recognition complex, as has been shown for the *Escherichia coli* UvrC protein^{10–14}. During mitotic recombination³, the ability of RAD10 to bind ssDNA and to renature complementary DNA strands could stimulate the activity of other proteins that catalyse the strand-exchange reaction. For instance, the T4 gene 32 protein stimulates the strand exchange reaction catalysed by the uvsX protein^{15–17}, the β protein of phage λ enhances the formation of joint molecules by the *E. coli* RecA protein¹⁸, and the yeast SF1 protein augments the strand-exchange reaction catalysed by the SEP1 protein¹⁹. The gene 32 protein^{20,21}, the β protein¹⁸, and the SF1 protein²² all bind ssDNA and promote the renaturation of cDNA strands. The yeast excision repair protein RAD1 functions with RAD10 in a mitotic recombination pathway that is distinct from the

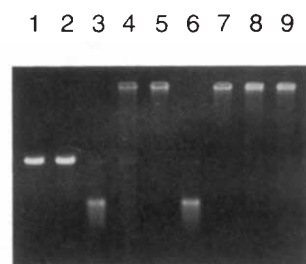


FIG. 3 RAD10 renatures complementary DNA strands. Agarose gel electrophoresis of renaturation products. All the samples had been treated with 1% SDS to release RAD10 from DNA before electrophoretic analysis. Lane 1, linear M13 dsDNA; lane 2, linear M13 dsDNA incubated with 1 μg RAD10. Lanes 3–9 contain complementary single DNA strands generated by heat-denaturation of linear M13 dsDNA. Lane 3, no RAD10; lane 4, 0.5 μg RAD10; lanes 5–9 each contains 1 μg RAD10. In lane 6, after incubation with RAD10, the reaction mixture was heated at 100 °C for 1 min before electrophoresis. ATP, Mg^{2+} , and both ATP and Mg^{2+} were omitted from lanes 7, 8 and 9, respectively. Addition of 1% SDS to the reaction mix at the beginning of incubation abolished the formation of high M_r DNA.

METHODS. Complementary single strands of DNA were generated from *Bgl*II linearized M13 mp18 DNA by boiling (2 min in H_2O) and immediate cooling on ice. Following the addition of reaction buffer (see Fig. 2) and RAD10 protein (0.5–1 μg) to 400 ng denatured DNA in a final volume of 10 μl , the reaction mixtures were incubated at 30 °C for 15 min. RAD10 protein was stripped from DNA by treatment with 1% SDS for 2 min at 30 °C before being analysed on 0.8% agarose gels electrophoresed in TAE buffer (40 mM Tris-acetate, pH 7.4, containing 2 mM EDTA). DNA bands were visualized by staining with ethidium bromide.

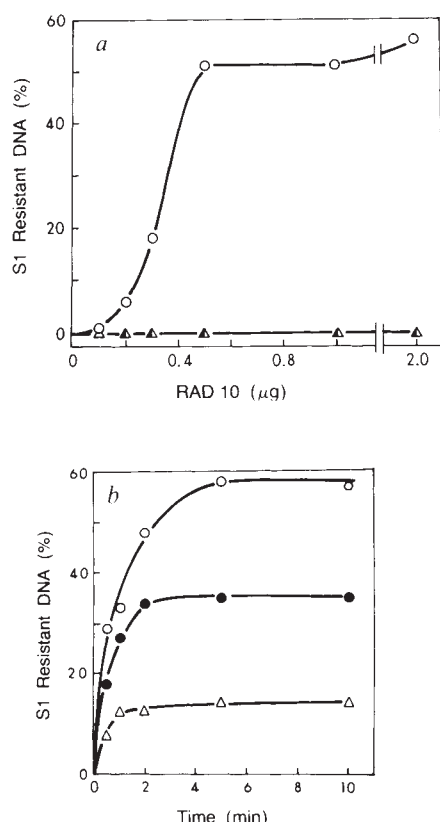


FIG. 4 Renaturation of DNA as determined by resistance to S1 nuclease digestion. *a*, Generation of S1 nuclease-resistant DNA requires homology. Renaturation of heat-denatured T7 DNA as a function of RAD10 concentration (○). Nonhomologous M13 viral ssDNA remained susceptible to S1 digestion after incubation with RAD10 (Δ). Addition of 1% SDS to the reaction mixture before addition of RAD10 abolished renaturation of heat denatured T7 DNA (▲). The incubation time for these experiments was 5 min. *b*, Kinetics of formation of S1-resistant DNA from heat-denatured T7 DNA. The amounts of RAD10 used were 300 ng (Δ), 400 ng (●) and 1 μg (○).

METHODS. ³H-labelled T7 DNA (400 ng) was heat-denatured (see Fig. 3) and incubated with RAD10 (0.1–2 μg) in 10 μl of reaction buffer (see Fig. 2) at 30 °C for 0.5–10 min. Omission of ATP and Mg²⁺ from the reaction mix had no effect on the formation of S1-resistant DNA. After the addition of 3 μl 10% SDS and a 3 min incubation at 37 °C, the volume of reaction mixtures was adjusted to 300 μl with S1 nuclease buffer (50 mM Na-acetate, pH 4.7, containing 150 mM NaCl and 1 mM Zn-acetate). S1 nuclease (50 units; BRL) was then added and digestion of ssDNA proceeded at 37 °C for 30 min. The reaction was terminated by the addition of 10 μl heat-denatured calf thymus DNA (10 mg ml⁻¹) and 1 ml 10% trichloroacetic acid (TCA). After 20 min on ice, precipitated DNA was collected by filtration through GF/C filters (Whatman), which were washed with 10 ml ice-cold 10% TCA and 7 ml ice-cold ethanol, dried, and mixed with 10 ml Ready Protein⁺ scintillation fluid (Beckman) before radioactivity measurement.

RAD52 recombination pathway^{3,23}. RAD10 could act with RAD1 in the strand-exchange reaction, or it may cooperate with another unidentified protein. □

11. Yeung, A. T., Mattes, W. B., Oh, E. Y. & Grossman, L. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6157–6161 (1983).
12. Yeung, A. T., Mattes, W. B. & Grossman, L. *Nucleic Acids Res.* **14**, 2567–2582 (1986).
13. Oh, E. Y. & Grossman, L. *Nucleic Acids Res.* **14**, 8557–8571 (1986).
14. Sancar, A. & Rupp, W. D. *Cell* **33**, 249–260 (1983).
15. Yonesaki, T. & Minagawa, T. *EMBO J.* **4**, 3321–3327 (1985).
16. Formosa, T. & Alberts, B. M. *J. biol. Chem.* **261**, 6107–6118 (1986).
17. Hinton, D. M. & Nossal, N. G. *J. biol. Chem.* **261**, 5663–5673 (1986).
18. Muniyappa, K. & Radding, C. M. *J. biol. Chem.* **261**, 7472–7478 (1986).
19. Norris, D. & Kolodner, R. *Biochemistry* **29**, 7903–7911 (1990).
20. Alberts, B. M. & Frey, L. *Nature* **227**, 1313–1318 (1970).
21. Kowalczykowski, S. C., Bear, D. G. & Von Hippel, P. H. *Enzymes* **14**, 373–444 (1981).
22. Norris, D. & Kolodner, R. *Biochemistry* **29**, 7911–7917 (1990).
23. Schiestl, R. H. & Prakash, S. *Molec. cell. Biol.* **8**, 3619–3626 (1988).
24. Sung, P., Prakash, L., Weber, S. & Prakash, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6045–6049 (1987).
25. McEntee, K., Weinstock, G. M. & Lehman, I. R. *J. biol. Chem.* **256**, 8835–8844 (1981).

ACKNOWLEDGEMENTS. We thank R. Kolodner and S. Kowalczykowski for radioactive DNAs and S. Chaudhury for the construction of plasmid pSUC8. This work was supported by the US National Cancer Institute and Department of Energy.

Selective inhibition of neurite outgrowth on mature astrocytes by Thy-1 glycoprotein

Marie-Catherine Tiveron*, Erminia Barboni†, F. Bernardo Pliego Rivero*, Ann Marie Gormley*, P. John Seeley*, Frank Grosveld‡ & Roger Morris*

* Norman and Sadie Lee Research Centre, Laboratory of Neurobiology, and ‡ Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

† Department of Human Biopathology, Facoltà di Medicina e Chirurgia, University of Rome 'La Sapienza', Rome, Italy

THY-1, the smallest member of the immunoglobulin superfamily¹, is a major cell-surface component expressed by several tissues². The protein³, carbohydrate³ and gene⁴ structures of this molecule are known, yet its function is not. It is highly expressed in nervous tissue⁵, where it appears on virtually all neurons after the cessation of axonal growth⁶. Here we show that expression of Thy-1 by a neural cell line inhibits neurite outgrowth on mature astrocytes, but not on other cellular substrata which include Schwann cells and embryonic glia. This inhibition of neurite extension on astrocytes can be reversed by low concentrations (nanomolar) of soluble Thy-1. If a similar interaction between neuronal Thy-1 and astrocytes occurs *in vivo*, it could stabilize neuronal connections and suppress axonal regrowth after injury in the astrocyte-rich areas of adult central nervous system.

The neural cell line NG115-401L (ref. 7) remains negative for Thy-1 (Thy-1⁻) as it extends bipolar neurites during differentiation. A clone (H1) expressing human Thy-1 and two clones (M1 and 2) expressing mouse Thy-1.2 (ref. 2) were obtained by transfection (see Fig. 1 legend (see page 747)); a unique site of insertion for each clone was verified by Southern blotting). Thy-1 was on both cell bodies and neurites (Fig. 1*r*). M1 and M2 clones, in which Thy-1 was driven by the strong β-actin promoter⁸, had a level of Thy-1 roughly equivalent to that of mature neurons² (by fluorescence-activated cell sorting, 85 ± 10% (M1) and 92 ± 12% (M2), using thymocytes as 100% control), whereas expression was lower for the H1 clone (4.8 ± 1.3%) levels. Three control clones (C1–3), and the parental NG115 cells, bound undetectable Thy-1 antibody (<0.1%). Differentiated cells of each Thy-1⁺ clone (prelabelled red with DiI (1,1 diiododecyl-3,3',3'-tetramethylindocarbocyanine perchlorate); ref. 9) were mixed with one of the Thy-1 clones (prelabelled green with DiO (3,3'-diiododecylsuccinylcarbocyanine perchlorate); ref. 9) and applied together to a dish on which a monolayer of test cells, covering 70% of the culture dish, were growing; the remaining plastic area provided an acellular control substrate.

Received 4 July; accepted 12 November 1991.

1. Wilcox, D. R. & Prakash, L. *J. Bact.* **148**, 618–623 (1981).
2. Reynolds, R. J. & Friedberg, E. C. *J. Bact.* **146**, 692–704 (1981).
3. Schiestl, R. H. & Prakash, S. *Molec. cell. Biol.* **10**, 2485–2491 (1990).
4. van Duin, M. et al. *Cell* **44**, 912–923 (1986).
5. Reynolds, P., Prakash, L., Dumais, D., Perozzi, G. & Prakash, S. *EMBO J.* **4**, 3549–3552 (1985).
6. Morandelli, F. & Lieberman, M. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3201–3205 (1980).
7. Anderson, R. A. & Coleman, J. E. *Biochemistry* **14**, 5485–5491 (1975).
8. McPherson, A., Jurnak, F. A., Wang, A. J., Molineux, I., & Rich, A. *J. molec. Biol.* **134**, 379–400 (1979).
9. Shimamura, A., Tremethick, D. & Worcel, A. *Molec. cell. Biol.* **8**, 4257–4269 (1988).
10. Sancar, A., Kacinski, B. M., Mott, D. L. & Rupp, W. D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5450–5454 (1981).

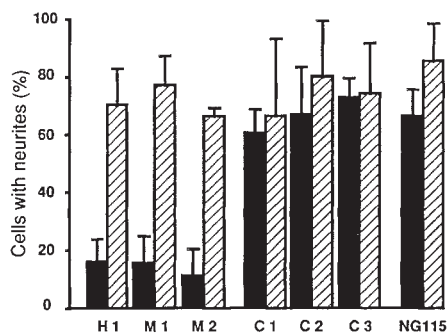


FIG. 2 Proportion of cells growing neurites upon astrocytes (filled bars) and plastic (hatched bars), mean $\pm$ s.d. Neurite outgrowth was assessed using phase contrast and fluorescence optics, 8 h after the cells were added to the astrocytes (similar results were obtained after 5–24 h of neurite outgrowth). Cells were counted as having a neurite if they extended a process longer than one cell diameter ($15\text{ }\mu\text{m}$); if the cell body or neurite at any stage extended over the cell monolayer, it was counted as a cell growing on the monolayer. Cells were inspected under both phase and fluorescent optics, 200–500 neural cells on both the monolayer and the plastic were counted. Thy-1⁺ clones growing on astrocytes differ from their growth on plastic, and from the Thy-1⁻ clones on either substrate; $P < 0.01$. The number of separate experiments were: H1, 20; M1, 7; M2, 4; C1, 7; C2, 5; C3, 5; NG115, 8. Data were analysed by Duncan's new multiple range test²⁵ on an Apple Macintosh using SuperAnova.

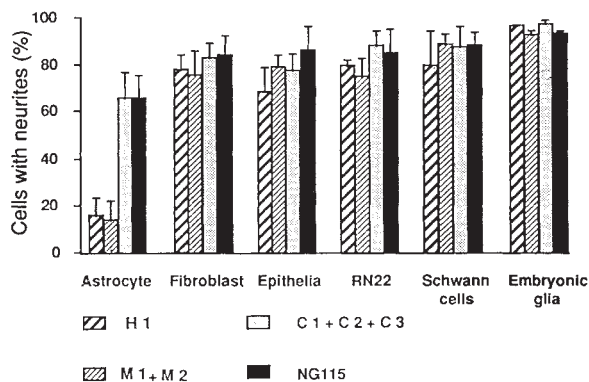


FIG. 3 Comparison of proportion of cells extending neurites over different cellular substrata, mean $\pm$ s.d. Dark stripe, H1 cells; fine stripe, M1 and 2 cells (pooled data); stipple, C1–3 (pooled data); solid bars, NG115 cells. Data were analysed as in Fig. 2. There was no significant difference between the M1 and M2 cells, or between any of the control (C) clones, so their data have been pooled. The Thy-1⁺ clones growing over astrocytes differ from all other groups ($P < 0.01$). The number of separate experiments were: 8–20 on astrocytes; 4–8 on fibroblasts; 3–8 on epithelia; 3–4 on RN22 schwannoma cells; 3–7 on Schwann cells; and 3 on embryonic glia (the smaller number in each case refers to control clones, the larger to the H1 and M clones). Astrocytes, prepared from 6-day-old rat corpus callosum²⁶ were 99.5% (at 1 month) to 95% (at 4 months) positive for glial fibrillary acidic protein and formed a flattened monolayer of cells; they were used after 2–5 months in culture. Schwann cells were prepared from 6-day-old rat sciatic nerve²⁷ and used within 4 days; RN22 schwannoma cells and DC1 kidney epithelial cells were cultured as described^{11,28}, meningeal fibroblasts were prepared from the meninges of 2-day-old rats²⁹, and used within 1–12 weeks; embryonic glia prepared (unpublished data) from the cortex of embryonic rats (16 days gestation) were used within 2 days. The Schwann cells and schwannoma cells promoted neurite outgrowth, in that the processes were distinctly longer ($100\text{--}150\text{ }\mu\text{m}$ for $\sim 20\%$ of the population) on these cells; on embryonic glia essentially all cells extended neurites, some in excess of $200\text{ }\mu\text{m}$. FACS analysis showed that the astrocytes and kidney epithelia had similar levels of surface Thy-1 ($6.3 \pm 1.4\%$ and $5.8 \pm 2.1\%$, respectively, the thymocyte level), fibroblasts had $1.9 \pm 0.8\%$ and RN 22 cells had $0.8 \pm 0.3\%$. Schwann and embryonic glial cells were Thy-1⁺.

A striking difference was found between all three Thy-1⁺ clones and the negative controls in their neurite outgrowth on astrocytes (Figs 1a–c and 2), but not on the adjacent plastic (Figs 1d and 2) nor on other cellular substrata, including the growth-promoting glial cells of embryonic brain and peripheral nervous system (Figs 1g–j and 3). On the astrocytes, about 75% of Thy-1⁺ cells extended bipolar neurites, each usually $40\text{--}100\text{ }\mu\text{m}$ (Figs 1a–c and 2). Most Thy-1⁺ cells, however, failed to send out any process, remaining a round ball attached to the astrocytes, often to the same cell over which Thy-1⁻ neurites had extended (Figs 1a–c and 2). On the plastic beside the astrocytes, both Thy-1⁺ and Thy-1⁻ cells elaborated neurites with equal vigour (Fig. 1d). Where a Thy-1⁺ neurite growing on plastic encountered the monolayer or a process emerging from it, the neurite invariably ended (Fig. 1e, f), whereas a Thy-1⁻ neurite usually extended over the cells. Those Thy-1⁺ cells counted as extending processes on astrocytes (Figs 2–4) usually had either a small ($15\text{--}20\text{ }\mu\text{m}$) neurite, or one of their long neurites growing on plastic extended a few microns onto the monolayer. The H1 clone, expressing 5% of the Thy-1 level of the M clones, showed equal suppression of neurite outgrowth on astrocytes.

These results imply that Thy-1 on the neural cell line interacts selectively with the astrocytes to limit neurite extension by the Thy-1⁺ cell. To test whether the astrocytic ligand could be blocked by purified Thy-1, we added the molecule along with the neural cells to astrocyte monolayers. Neurite outgrowth by the Thy-1⁺ cells was restored in a dose-dependent manner. At 200 nM Thy-1, the proportion of cells extending neurites (Fig.

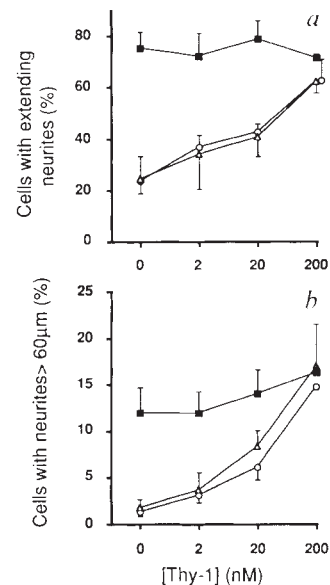
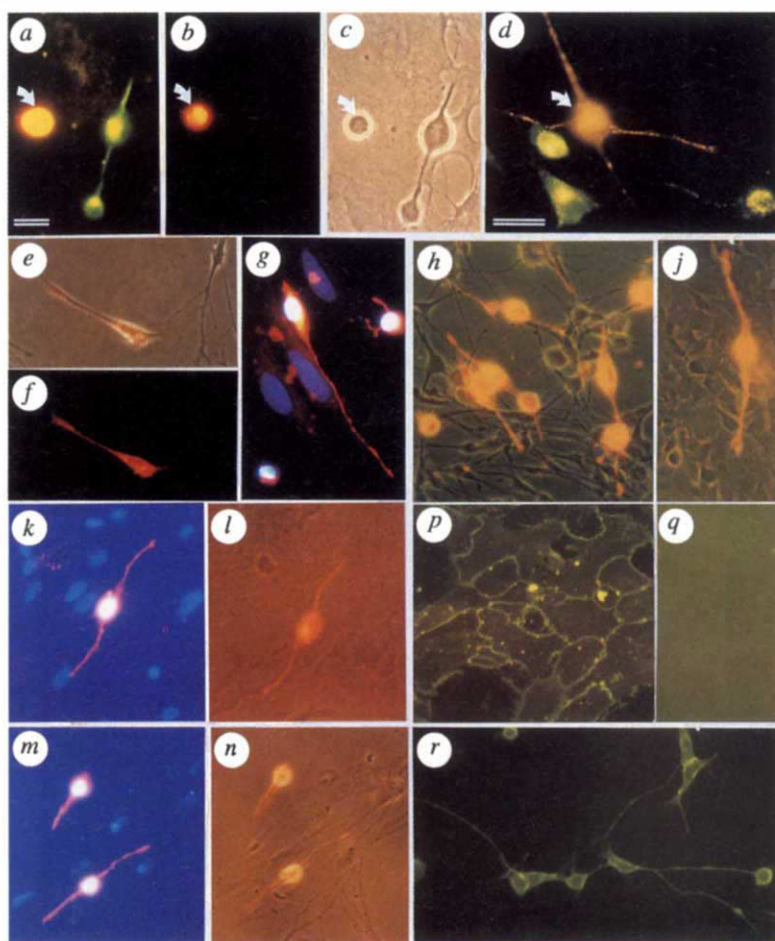


FIG. 4 Pure Thy-1 enables Thy-1⁺ cells to extend neurites on astrocytes. a, Proportion of NG115 (■), H1 (○) and M1 (△) cells extending neurites over astrocytes in the presence of pure Thy-1 (mean $\pm$ s.d.; $n=3$ (NG115 and H1) and 5 (M1)). b, Proportion of cells (as in a) extending a neurite longer than $60\text{ }\mu\text{m}$ over astrocytes.

METHODS. Thy-1, prepared from rat brain using immunoaffinity chromatography⁶, was dialysed against 1mM HEPES, pH 8.5, freeze-dried, washed 3 times with absolute ethanol to remove deoxycholate¹¹, resuspended in DMEM, and centrifuged at $14,000\text{ r.p.m.}$ for 15 min in a microfuge before addition to the cells. Such detergent-free Thy-1 forms a 16-mer micelle, internalizing the phosphatidylinositol tail^{1,30}. Neurite outgrowth over astrocytes was assayed as in Fig. 2. Thy-1 concentration was determined by radioimmunoassay using ¹²⁵I-labelled OX7 Fab anti-Thy-1.1, and rat thymocytes as a standard¹¹. The concentration shown is of the Thy-1 molecule, not the micelle. Eluate from the immunoaffinity column, taken after the pool of Thy-1 was collected, was similarly dialysed, freeze-dried and extracted with ethanol, then washed with DMEM. This medium had no effect on neurite outgrowth by any of the cells (results not shown).

FIG. 1 Neurite outgrowth on different substrata by Thy-1⁺ (H1, M1) and Thy-1⁻ (C1, C2) cells. *a-c*, H1 (arrow, labelled red) and C1 (green) cells on astrocytes, seen with green (*a*) and red fluorescence (*b*), and phase optics (*c*). *d*, Cells growing on plastic of the same dish (arrow, H1). *e, f*, H1 cell on plastic, its short neurite touches but does not cross a fine branch of an astrocytic process. *g*, M1 cell growing neurite over embryonic glia (nuclei visible with Hoechst dye). *h* and *j*, M1 cells growing over Schwann (*h*) and kidney epithelial (*j*) cells (combined phase and fluorescence). *k* and *l*, M1 cells growing over astrocytes in the presence of 200 nM Thy-1. *m* and *n*, C2 cells (here labelled red) with 200 nM Thy-1 on astrocytes and plastic. *p* and *q*, CBA mouse astrocytes incubated with 200 nM rat Thy-1 (*p*) or Thy-1-negative column eluant (*q*; Fig. 3 legend) for 1 h at 37 °C, washed 3 times and immunolabelled for rat Thy-1. *r*, M1 cells on plastic, immunolabelled for Thy-1.2. Scale bars in *a* and *d* (rest at same magnification as in *a*) are 20 µm.

METHODS. Clones expressing human (H1) and mouse (M1, M2) Thy-1 were obtained by transfection¹² of NG115-401L cells with human Thy-1 under the mouse Thy-1 promoter (mhm construct⁴), or mouse Thy-1 (Thy-1.2 allele²) regulated by the human β -actin promoter (1.6 kb upstream of the *Pst*I site in the first Thy-1 exon was replaced by a 4.3-kb fragment of pHBAPr-1-neo⁸). Three Thy-1⁻, neomycin-resistant clones were retained as controls (C1-3). Cells, differentiated for 4 days in DMEM medium plus 1 mM pyruvate, 1% FCS, 1 mM dbcAMP (dibutyl cAMP) and 1.5% dimethylsulphoxide, followed by 24–36 h in the same plus 10 µM retinoic acid, 100 nM insulin, 100 nM dexamethasone and 50 ng ml⁻¹ 2.5 nerve growth factor (NGF; sedimentation at 2.5S) (S.I. Patterson, W.T. Cheung and M.R. Hanley, personal communication) were pre-labelled with Dil or DiO⁹, suspended by trypsinization, washed and overlain on test cells. After 8 h in differentiation medium, plates were washed twice with PBS, fixed with 1% paraformaldehyde and the nuclei labelled with Hoechst dye.



3a) and the proportion with longer neurites (see, for example, Figs 1*k, l* and 3*b*) was equivalent to that of the parental NG115 cells. Binding of exogenous 200 nM Thy-1 to astrocytes was observed (Fig. 1*p, q*); this was still detectable by immunofluorescence at 20 nM, but not 2 nM, Thy-1 (data not shown). This indicates an affinity for the interaction well within the range observed for other cell-surface members of the immunoglobulin superfamily¹⁰.

Direct blocking of Thy-1 on the neural cell lines by pre-incubating them with 150 µg ml⁻¹ of rabbit F(ab)₂ anti-Thy-1 antibodies¹¹ also restored Thy-1⁺ neurite outgrowth on the astrocytes: from 10.1 ± 1.9% (M1) and 8.6 ± 4.8% (H1) with non-immune F(ab)₂, to 59.8 ± 5.7% (M1) and 63.1 ± 11.8% (H1) with antibody (*n* = 3, mean ± s.d.).

The astrocytic ligand for Thy-1 is unlikely to be Thy-1 itself because it is expressed on the kidney epithelial cells and fibroblasts (see Fig. 4 legend) which did not inhibit Thy-1⁺ neurite outgrowth. We have not yet identified when this ligand appears on astrocytes, although their inhibitory effect on Thy-1⁺ neurite extension becomes apparent after 1–2 months in culture (our unpublished results).

The action of Thy-1 could be indirect, with activation of the Thy-1 ligand inducing the astrocyte to modify its surface so as to inhibit neurite extension. This seems improbable, because a Thy-1⁻ cell often grew a neurite over the same astrocyte to which a rounded Thy-1⁺ cell was attached (for example, Fig. 1*a-c*); nor does pure Thy-1 inhibit neurite outgrowth by Thy-1⁻ cells (Figs 1*m, n* and 3). We therefore suggest that Thy-1 directly initiates a neurite growth-inhibitory signal across the membrane of the neural cell, presumably by a mechanism similar to its action in T-cell signalling^{12,13}.

So does neuronal Thy-1 downregulate the growth of axons

contacting mature astrocytes? Its appearance in development on most neurons towards the end of axonal growth⁶ is compatible with this role. Thy-1 is found much earlier in development on sympathetic neurons¹⁴, whose axons are unusual in being restricted to the peripheral nervous system and so do not encounter astrocytes. Moreover, Thy-1 is never found, even in the adult, on olfactory neurons², whose axons are unique in being able to grow into adult central nervous system¹⁵. Thus Thy-1 appears to be excluded from axons while they are able to grow in the central nervous system.

The influence of astrocytes on axonal growth, however, is at present unclear, and undoubtedly complex. It has frequently been suggested on the basis of *in vivo* experiments that astrocytes restrict axonal growth^{16–20}, although they might simply fail to provide a stimulation essential to growth. In tissue culture, astrocytes either permit or inhibit^{21–24} neurite growth from primary cultured neurons, depending on the state of differentiation of both astrocytes^{22,23} and neurons²⁴, and the spatial configuration of the astrocytes²⁴. Further progress requires the role of individual molecules involved in the neuron-astrocyte interaction to be defined, and here Thy-1 and its ligand are prime candidates. □

Received 7 August; accepted 2 December 1991.

- Williams, A. F. in *Cell Surface Thy-1* (eds Reif, A. E. & Schlesinger, M.) 49–69 (Dekker, New York, 1989).
- Morris, R. J. *Dev. Neurosci.* **7**, 133–160 (1985).
- Parekh, R. B., Tse, A. G. D., Dwek, R. A., Williams, A. F. & Rademacher, T. W. *EMBO J.* **6**, 1233–1244 (1987).
- Vidal, M., Morris, R., Grosfeld, F. & Spanopoulou, E. *EMBO J.* **9**, 833–840 (1990).
- Reif, A. E. & Allen, J. M. V. *J. exp. Med.* **120**, 413–433 (1964).
- Xue, G. P., Pliego Rivero, F. B. & Morris, R. J. *J. Development* **112**, 161–176 (1991).
- Hatanaka, H. & Amano, H. *Brain Res.* **215**, 305–316 (1981).

8. Gunning, P., Leavitt, J., Muscat, G., Ng, S.-Y. & Kedes, L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4831–4835 (1987).
9. Honig, M. G. & Hume, R. I. *J. Cell Biol.* **103**, 171–187 (1986).
10. Springer, T. A. *Nature* **353**, 704–705 (1991).
11. Morris, R. J., Mancini, P. E. & Pfeiffer, S. E. *Brain Res.* **182**, 119–135 (1980).
12. Barboni, E., Gormley, A. M., Pliego Rivero, F. B., Vidal, M. & Morris, R. J. *Immunology* **72**, 457–463 (1991).
13. Robinson, P. J. *Immun. Today* **12**, 35–41 (1991).
14. Birren, S. J. & Anderson, D. J. *Neuron* **4**, 189–201 (1990).
15. Graziadei, P. P. C. & Monti Graziadei, G. A. in *Neuronal Plasticity* (ed. Cotman, C. W.) 131–153 (Raven, New York, 1978).
16. Ramon y Cajal, S. *Degeneration and Regeneration of the Nervous System* Vol. II (Hafner, New York, 1927).
17. Carlstedt, T., Cullheim, S., Risling, M. & Ulfhake, B. *Brain Res. Bull.* **22**, 93–102 (1989).
18. Liuzzi, F. J. & Lasek, R. J. *Science* **237**, 642–645 (1987).
19. Siegal, J. D., Klot, M., Smith, G. M. & Silver, J. *Expl. Neurol.* **109**, 111–130 (1990).
20. Stensaas, L. J., Partlow, L. M., Burgess, P. R. & Horsch, K. W. *Progr. Brain Res.* **71**, 457–468 (1987).
21. Noble, M., Fok-Seang, J. & Cohen, J. J. *Neurosci.* **4**, 1892–1903 (1984).
22. Smith, G. M., Rutishauser, U., Silver, J. & Miller, R. H. *Dev. Biol.* **138**, 377–390 (1990).
23. Baehr, M. & Bunge, M. *Glia* **3**, 293–300 (1990).
24. Fawcett, J. W., Housden, E., Smith-Thomas, L. & Meyer, R. L. *Dev. Biol.* **135**, 449–458 (1989).
25. Duncan, D. B. *Biometrics* **11**, 1–42 (1955).
26. Emmett, C. J., Lawrence, J. M. & Seeley, P. J. *Brain Res.* **447**, 223–233 (1989).
27. Brookes, J. P., Fields, K. L. & Raff, M. C. *Brain Res.* **165**, 105–118 (1979).
28. Kollias, G. *et al. Cell* **51**, 21–31 (1987).
29. Raff, M. C. *et al. Brain Res.* **174**, 283–308 (1979).
30. Kuchel, P. W., Campbell, D. G., Barclay, A. N. & Williams, A. F. *Biochem. J.* **169**, 411–417 (1978).

ACKNOWLEDGEMENTS. We thank M. Hanley for the NG115–401L cells and the protocol for their differentiation, J. Lawrence for help with the Schwann cell cultures, J. Beech for technical assistance, I. Robinson and T. Wells for help with statistics, and G. Raisman and C. Clemente for discussions. M.C.T. and A.M.G. were supported by the International Spinal Research Trust, F.B.P.R. by the British Neurological Research Trust. This work was supported by the MRC and the above Trusts.

Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli*

Penelope E. Stein*, Amechand Boodhoo*,
Gregory J. Tyrrell†, James L. Brunton‡
& Randy J. Read*§

* Department of Medical Microbiology and Infectious Diseases,
University of Alberta, Edmonton, Alberta T6G 2H7, Canada

† Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto,
Ontario M5G 1X5, Canada;

and Departments of Microbiology and ‡ Medicine, University of Toronto,
Toronto, Ontario M5S 1A1, Canada

THE Shiga toxin family, a group of cytotoxins associated with diarrhoeal diseases and the haemolytic uraemic syndrome, includes Shiga toxin from *Shigella dysenteriae* type 1 and verotoxins¹ produced by enteropathogenic *Escherichia coli*. The family belongs to the A–B class of bacterial toxins, which includes the cholera toxin family, pertussis and diphtheria toxins. These toxins all have bipartite structures consisting of an enzymatic A subunit associated with a B oligomer which binds to specific cell-surface receptors, but their amino-acid sequences and pathogenic mechanisms differ. We have determined the crystal structure of the B oligomer of verotoxin-1 from *E. coli*. The structure unexpectedly resembles that of the B oligomer of the cholera toxin-like heat-labile enterotoxin from *E. coli*², despite the absence of detectable sequence similarity between these two proteins. This result implies a distant evolutionary relationship between the Shiga toxin and cholera toxin families. We suggest that the cell surface receptor-binding site lies in a cleft between adjacent subunits of the B pentamer, providing a potential target for drugs and vaccines to prevent toxin binding and effect.

The crystal structure determination of the B oligomer of verotoxin 1 (VT-1) is described in Table 1. The fold is very similar to that of the B oligomer of heat-labile enterotoxin from *E. coli* (LT; ref. 2). The VT-1 B subunit structure (Fig. 1) is a pentamer in which each monomer (69 amino acids) comprises two three-stranded antiparallel β -sheets and an α -helix. As in LT, β -sheets from pairs of adjacent monomers form six-stranded

antiparallel β -sheets at the outer surface of the pentamer, while the five helices line a 'pore' in the centre. The pore has a diameter of about 11 Å and is lined by neutral and non-polar residues, contrasting with the highly charged pore of LT².

The structural similarity between the B pentamers of VT-1 and LT is surprising, given the difference in size (69 amino acids for VT-1, 103 for LT), and toxic action of the associated A subunit (VT-1 is an *N*-glycosidase, LT is an ADP-ribosyl-

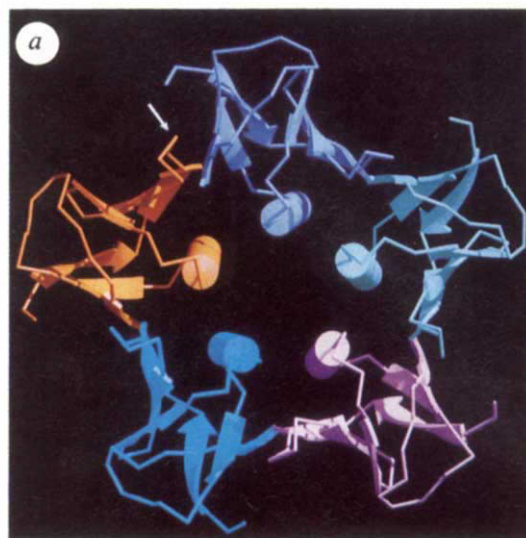


FIG. 1 Schematic drawings of the structure of the VT-1 B pentamer viewed along (a) and perpendicular to (b) the 5-fold axis. β -Strands are represented by arrows and α -helices by cylinders. Secondary structure elements calculated with the program DSSP²³ are: β 1, 3–8; β 2, 9–14; β 3, 20–24; β 4, 27–31; α , 36–46; β 5, 49–53; β 6, 65–68. There is a β -sheet interaction between β 2 of monomer N and β 6 of monomer N+1 (numbering anti-clockwise when viewed as in a with monomer 1 at the top), except between monomers 1 and 2 (arrow) where these strands are separated by about 5 Å. This gap results from a screw component of about 1.3 Å in the 5-fold rotation axis. Four pairs of monomers are related by a rotation of approximately 72° plus a translation of 1.3 Å along the rotation axis, so that the last pair of monomers does not align. The screw component may explain why the fivefold symmetry was not evident in the self-rotation function (ref. 16, and our unpublished results). In LT, by contrast, the pentamer is nearly perfect². The VT-1 pentamer is much less stable to disruption by SDS than is the LT pentamer²⁴. It is conceivable that instability of the pentamer might be relevant to the mechanism of A-subunit translocation across the cell membrane, but considering the close structural similarity to LT, a special mechanism for VT-1 seems unlikely. We suspect, therefore, that the lack of symmetry is an artefact of crystallization. A zinc atom is bound between the N-termini of two non-equivalent B monomers (2 and 4) from crystallographically related pentamers, consistent with the improvement in crystal morphology when zinc was included in the crystallisation medium. The crystal contact between monomers 2 and 4 is by far the most extensive, and may be related to the disruption of the pentamer between monomers 1 and 2.

§ To whom correspondence should be addressed.

We tried to locate the carbohydrate-binding site in the VT-1 B subunit structure by analysing the aligned amino-acid sequences of the B subunits of members of the Shiga toxin family (residue numbering used is based on the VT-1 sequence). The sequences of Shiga toxin, verotoxin-2 (VT-2) and pig oedema

The cleft (Figs 2b and 3) has many characteristics typical of carbohydrate-binding sites in proteins⁹. It is occupied by

TABLE 1 Summary of native and heavy atom derivative data

	Native	CdCl ₂	K ₂ PtCl ₄	Trimethyl lead acetate	AgNO ₃	Mersalyl acid
Number of crystals	1	2	1	1	1	1
Total reflections	84,281	106,731	75,124	62,713	72,494	86,978
Complete (%)	98	98	85	85	71	93
R _{merge} (%) [*]	5.4	5.5	6.2	4.8	5.2	4.6
R _{deriv} (%) [†]	—	17.2	33.8	19.8	15.0	13.7
Number of sites	—	2	7	4	2	3
Phasing power [‡]	—	0.46	1.24	1.14	1.05	0.65

VT-1 was purified and crystallized as previously described by us¹⁵ and independently by Hart *et al.*¹⁶, except that addition of 1.5 mM zinc chloride to the crystallization medium increased the thickness of the crystals and improved their morphology. The crystals belong to space group P2₁2₁2₁ (cell dimensions $a=59.6$ Å, $b=102.4$ Å, $c=56.1$ Å) with one B-subunit pentamer per asymmetric unit. Data to 2.2 Å from native crystals (co-crystallized with zinc) and from five heavy-atom derivatives were collected using a San Diego-style multiwire counter area detector with rotating anode source. Heavy-atom sites were found using difference Patterson and difference Fourier analyses. Heavy-atom parameters were refined with a modified version of the program PHARE¹⁷ resulting in an overall figure of merit of 0.58 (25–2.5 Å). The amino-acid sequence of one B subunit monomer was fitted to an electron-density map calculated with unmodified multiple isomorphous replacement (MIR) phases using the program FRODO¹⁸, based on a skeletonized map produced with the program BONES¹⁹. The helical sequence (residues 36–46) was fitted in all five B monomers and the symmetry operations that superimposed these helices were used to generate the complete B pentamer. The model was improved by rigid-body refinement followed by simulated annealing with molecular dynamics using the program XPLOR²⁰ and then by alternate rounds of restrained parameter least-squares refinement with the program PROLSQ²¹ and model building with FRODO. Model phases were combined with the original MIR phases using the program SIGMAA²² during the initial rounds of refinement. The current model, which includes 2,700 protein atoms, 102 waters and 3 zinc atoms, has a crystallographic R-factor of 17.7% for all data between 10 and 2.2 Å. The r.m.s. deviation from ideal values for bond lengths is 0.017 Å and for bond angles is 2.9°. The r.m.s. deviation of atomic coordinates of main chain atoms is about 0.3 Å between monomer pairs 1–3, 1–4, 1–5, 3–4, 3–5 and 4–5, but about 0.4 Å between monomer pairs 2–1, 2–3, 2–4 and 2–5 (monomer numbering is described in the Fig. 1 legend).

$$^* R_{\text{merge}} = \sum_i |I_i - \langle I \rangle| / \sum_i I_i$$

$$^{\dagger} R_{\text{deriv}} = \sum_i |F_{\text{PH}}| - |F_{\text{P}}| / \sum_i |F_{\text{P}}|$$

$$^{\ddagger} \text{Phasing power} = [\sum_i |F_{\text{H}}|^2 / \sum_i (|F_{\text{PH}}(\text{obs})| - |F_{\text{PH}}(\text{calc})|)^2]^{1/2}$$

residues with polar and acidic side chains which could make hydrogen-bond interactions with polar groups of the carbohydrate, and contains exposed aromatic residues which could stack against the sugar rings. Site-directed mutagenesis has shown that at least two amino acids lining this cleft, Asp 17 and Lys 53, are involved in carbohydrate binding to Shiga toxin¹⁰. By contrast, loss of Gb₃ binding in two VT-2 mutants, Ala 43 → Thr and Gly 60 → Asp (ref. 11) is probably explained by poor folding of the protein; Ala 43 lies on a buried surface of the helix, and Gly 60 has a left-handed α -helical conformation in a turn. The proposed carbohydrate-binding site provides a structural explanation for the altered binding specificity associated with the double mutation Gln 65 → Glu, Lys 67 → Gln in VT-E (ref. 12). The mutant protein showed selective loss of Gb₄ binding with retained Gb₃ binding, consistent with binding of the terminal sugar residue of Gb₄ to the side chains of amino acids 65 and 67. These amino acids are found in the lower portion of the proposed carbohydrate-binding site (as it is shown in Fig. 2), suggesting that the glycolipid binds with the ceramide end (hence the cell membrane) at the upper extreme of the binding cleft. If the interaction between the A and B subunits of VT-1 is the same as in LT (ref. 2), with the A subunit lying below the B pentamer as it is shown in Fig. 2, the holotoxin would bind to cells with the A subunit away from the membrane. This agrees with the proposed membrane-binding orientation of LT and cholera toxin, based on the recently determined crystal structure of LT complexed with lactose¹³. □

- Seidah, N. G. *et al.* *J. Biol. Chem.* **261**, 13928–13931 (1986).
- Jaciewicz, M., Clausen, H., Nudelman, E., Donohue-Rolfe, A. & Keusch, G. T. *J. exp. Med.* **163**, 1391–1404 (1986).
- Lingwood, C. A. *et al.* *J. Biol. Chem.* **262**, 8834–8839 (1987).
- Waddell, T., Head, S., Petric, M., Cohen, A. & Lingwood, C. *Biochem. biophys. Res. Commun.* **152**, 674–679 (1988).
- DeGrandis, S., Law, H., Brunton, J., Gyles, C. & Lingwood, C. A. *J. Biol. Chem.* **264**, 12520–12525 (1989).
- Lindberg, A. A. *et al.* *J. Biol. Chem.* **262**, 1779–1785 (1987).
- Quiocho, F. A. *Pure appl. Chem.* **61**, 1293–1306 (1989).
- Jackson, M. P., Wadolkowski, E. A., Weinstein, D. L., Holmes, R. K. & O'Brien, A. D. *Bacteriology* **172**, 653–658 (1990).
- Perera, L. P., Samuel, J. E., Holmes, R. K. & O'Brien, D. J. *Bact.* **173**, 1151–1160 (1991).
- Tyrrell, G. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Sixma, T. K. *et al.* *Nature* **355**, 561–564 (1992).
- Bernstein, F. C. *et al.* *J. molec. Biol.* **112**, 535–542 (1977).
- Booth, A., Read, R. J. & Brunton, J. *J. molec. Biol.* **221**, 729–731 (1991).
- Hart, P. J. *et al.* *J. molec. Biol.* **218**, 691–694 (1991).
- Otwinski, Z. *Am. crystallogr. Ass. 1990 Annual Meeting, New Orleans Abstr.* C04 (1990).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268–272 (1978).
- Jones, T. A. & Thirup, S. *EMBO J.* **5**, 819–822 (1986).
- Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
- Hendrickson, W. A. & Konert, J. H. In *Computing in Crystallography* (eds Diamond, R., Ramaseshan, S. & Verkatesan, K.) 13.01–13.23 (Indian Institute of Sciences, Bangalore, 1980).
- Read, R. J. *Acta crystallogr.* **A42**, 140–149 (1986).
- Kabsch, W. & Sander, C. *Biopolymers*, **22**, 2577–2637 (1983).
- Ramotar, K. *et al.* *Biochem. J.* **273**, 805–811 (1990).

ACKNOWLEDGEMENTS. This work was supported by the MRC of Canada (J.L.B., R.J.R. and P.E.S.), the Alberta Heritage Foundation for Medical Research, the Ontario Ministry of Health (G.J.T.) and by an award to R.J.R. from the Howard Hughes Medical Institute. The coordinates of the VT-1 B subunit have been deposited with the Brookhaven Protein Data Bank¹⁴ (entry name, 1BOV).

ERRATUM

Mouse *Small eye* results from mutations in a paired-like homeobox-containing gene

Robert E. Hill, Jack Favor, Brigid L. M. Hogan, Carl C. T. Ton, Grady F. Saunders, Isabel M. Hanson, Jane Prosser, Tim Jordan, Nicholas D. Hastie & Veronica van Heyningen

Nature **354**, 522–525 (1991).

As a result of an error in the *Nature* office, an uncorrected version of this letter appeared in the 19/26 December issue. In the legend to Fig. 3 and in Figs 3b, 3d and 4 the designation *Sey*^{IN} should read *Sey*^{Neu} as in the text, and a revised Fig. 3e is shown below.

The penultimate sentence in paragraph 2 should read “In the homozygous *Sey/Sey* embryo (Fig. 1c) no eyes are visible and the nasal cavities do not develop^{15,16} resulting in narrowed, imperforate snouts.” The penultimate sentence in paragraph 3 should read: “As the mouse homologue of the WT-1 (Wilm’s tumour) gene is also deleted (data not shown), the *Sey*^H deletion covers a large region.”

